国外电子与通信教材系列

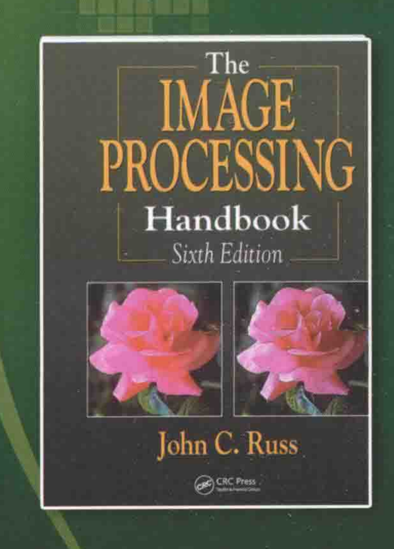


数字图像处理

（第六版）

The Image Processing Handbook

Sixth Edition



［美］ John C. Russ　著

余翔宇　等译

電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn

数字图像处理（第六版）

The Image Processing Handbook, Sixth Edition

本书是作者50年来在科研、教育和应用领域经验的总结，适合于读者自学，是数字图像处理领域的经典参考图书。全书采用实例来讲解图像处理的基本原理，不要求读者具备高深的数学知识，因此可让读者更好地了解应用各种处理技术的方式和原因。内容涉及图像增强、彩色图像处理、二维图像和三维结构间的关系等，比较了处理操作和测量操作的不同，涵盖了从显微学到天文学的宽广范围，并紧密结合了真实的科学应用、工业应用和前沿应用。

对于由显微镜、空间探测器和人眼获取的图像，都可使用相同的工具来获取、处理与分析图像中的数据。本书的特色如下：

- 按照从获取到解释的顺序，给出了各种图像处理与分析的方法，并对不同方法进行了比较。
- 介绍了改善图像外观、打印效果和传送效率的方法。
- 讲解了准备图像来测量特征和结构的过程。
- 详述了隔离结构进而测量其大小、形状、颜色和位置的方法。
- 介绍了校正图像缺陷与各种限制的方法。
- 说明了增强可视内容和解释细节的方法。
- 采用体视学方法介绍了测量结构的方法。
- 应用统计方法进行分类和识别。

作者简介

John C. Russ 美国北卡罗来纳州立大学教授，美国显微镜学会、微束分析学会、国际光学工程学会、国际体视学学会会员，定量形态学学会主席团成员，英国皇家显微学会会士。

说明：本书采用黑白方式印刷。请登录华信教育资源网（www.hxedu.com.cn），注册后免费下载原书中所有彩图。

其他相关图书
请扫二维码↓







定价：118.00 元



策划编辑：谭海平
责任编辑：谭海平
责任美编：孙焱津



国外电子与通信教材系列

数字图像处理

（第六版）

The Image Processing Handbook
Sixth Edition

［美］ John C. Russ 著
余翔宇 等译

電子工業出版社
Publishing House of Electronics Industry

内 容 简 介

这是一本关于图像处理基本原理及其具体应用的图书，是作者多年来教学、科研与应用的总结。全书共15章，内容贯穿从图像获取到应用的整个过程，具体包括图像获取、人类视觉、打印和存储、成像缺陷修正、空间域图像增强、频率域图像处理、分割和阈值处理、二值图像处理、全局图像测量、特定特征的测量、形状表征、特征识别与分类、层析成像、三维可视化、表面成像。本书的特色是，着重于各种图像处理方法的介绍与比较，并使用实例进行演示与说明。

本书的读者对象是信号与信息处理、通信工程、电子科学与技术、信息工程、自动化、计算机科学与技术、地球物理、生物工程、生物医学工程、物理、化学、医学、遥感等领域的大学教师和科技工作者、研究生、大学本科高年级学生及工程技术人员。

The Image Processing Handbook, Sixth Edition, John C. Russ
ISBN: 978-0-1323-4563-7

图书在版编目（CIP）数据

数字图像处理：第 6 版/（美）鲁斯（Russ J. C.）著；余翔宇等译. —北京：电子工业出版社，2014.8
书名原文：The Image Processing Handbook, Sixth Edition
国外电子与通信教材系列
ISBN 978-7-121-21439-4
I. ①图…　II. ①鲁…　②余…　III. ①数字图像处理－高等学校－教材　IV. ①TN911.73
中国版本图书馆 CIP 数据核字（2014）第 156369 号

策划编辑：谭海平
责任编辑：谭海平　特约编辑：王　崧
印　　刷：三河市鑫金马印装有限公司
装　　订：三河市鑫金马印装有限公司
出版发行：电子工业出版社
　　　　　北京市海淀区万寿路 173 信箱　邮编　100036
开　　本：787×1 092　1/16　印张：36.25　字数：1073 千字
版　　次：2014 年 8 月第 1 版（原著第 6 版）
印　　次：2014 年 8 月第 1 次印刷
定　　价：118.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888。
质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。
服务热线：（010）88258888。

译 者 序

图像处理是对图像信息进行加工以满足人的视觉心理或应用需求的行为。图像作为人类感知世界的视觉基础，是人类获取、表达和传递信息的重要手段。

John C. Russ 教授所著的本书在美国已出版了六版，受到了读者的一致肯定。尽管国内已引进了 Castleman、Gonzalez、Pratt 等学者的图像处理书籍，但本书几乎不用数学公式而尽量用文字和图例来解释图像处理的表达手法，使得其具有非常独特的特点。应电子工业出版社的邀请，本人有幸将该书翻译成中文，供广大国内读者参考。

本书的翻译延续了本人其他译著的风格。对于翻译过程中遇到的疑似错误，均以电子邮件的形式与原作者进行了沟通，并得到了作者的及时反馈。由于书中公式少，文字多，涉及的知识跨越多个学科，翻译难度极大；为力求准确，对于不常见的词汇，均采用了以英文单词在中国期刊网搜索关键词的方式进行初译，再以中文译名在谷歌等搜索引擎中搜索的方式进行确定。其他一些词汇则通过查阅广州图书馆及华南理工大学图书馆馆藏的相关书籍进行确定。对于实在无法查到对应中文表达的，则按照本人的理解给出翻译。即便如此，相信翻译的错漏仍在所难免，希望广大读者能够及时指出，以便在之后的重印中纠正。

本书第 3 章～第 5 章及第 9 章～第 14 章由本人独立翻译，其他章节由他人协助本人完成初译，再由本人整理完成，其中谢元杰和余伯庸参与第 1 章、第 6 章及第 8 章的初译和后期的校对整理，参与第 1 章初译的还有刘宁昊，参与第 2 章初译的有岑宣良、蓝俊健、黄伟鹏、张志欣及郑显华，参与第 6 章初译的还有张珂，参与第 7 章初译的有岑宣良、苏茜菲、孙祎楠和陈旭升，参与第 8 章初译的还有郭宇隆和朱祖聪，赵恒煜协助进行了第 15 章的初译。华南理工大学电子与信息学院杨春玲教授、胡永健教授、王伟凝副教授、周智恒副教授及刘琲贝老师，厦门大学的曾德炉副教授审阅了本书的译稿。梁宇文同学试读了本书的部分章节并提出了宝贵的意见。在此一并表示感谢。

译　者

致　谢

本书中包含的所有图像处理及结果图形的创建，均是在苹果计算机及索尼 VAIO 计算机上使用 Adobe Photoshop CS4 和 Fovea Pro 插件进行的。许多图像直接取自显微镜和直接向计算机提供数字输出的其他设备。其他图像则是用各种数码相机（如索尼、尼康、佳能等）获得的，有些图像是使用平板和幻灯片扫描仪（尼康和爱普生）获取的，大部分图像由我的同事和其他研究人员提供。无论这些图像的源头是否能确定，都在此表示感谢。

本书是以数字形式（存储在一张可擦写 DVD 上）提交给出版商的，并没有中间的打印稿及图像的底片或正片等。这意味着作者必须对排字错误或图形问题负全部责任。我已尽最大努力以像素级细节来显示放大的重要图像。这一过程也迫使我去了解一些出版技术。无论如何，从磁盘文件直接成书在确保了图像质量的同时，也缩短了图书制作的时间并降低了成本。衷心感谢 CRC 出版公司的编辑对我的培训并容纳书中数量众多的插图[①]（超过 2000 幅图形和 25 万个单词）。

特别感谢北卡罗来纳州 Reindeer Graphics 公司的 Chris Russ 先生，他对许多算法进行了编程并给出了宝贵的建议；还要特别感谢 Helen Adams 女士，她校对了本书中的许多内容，忍受了大量关于有效表达信息方式的讨论，并为此书的撰写计划得以实现提供了许多支持。

John C. Russ

于美国北卡罗来纳州罗利市

① 为节省读者购书成本，本书采用黑白印刷方式。原书中所有彩图，请至华信教育资源网（www.hxedu.com.cn）注册后，免费下载。——编者注

序　言

各种应用中因以下两个稍微不同的目的使用了图像处理：

⑴改进呈现给人类观看者的图像的外观，包括图像的印刷与传输。

⑵提供图像，以测量其揭示的特征与结构。

适用于每种任务的技术并非总是相同的，但二者有很大的交叠性。本书将讨论对两种任务都适用的方法。

为使工作尽可能最佳，知道所处理的图像将用于何处是很重要的。对于视觉增强而言，这意味着要在一定程度上熟悉人类视觉过程，并了解观察者对图像中的哪些线索有反应。关于人类视觉的一章涉及了这些问题。由于许多图像是在复制或传输的情况下进行处理的，因此了解图像的印刷或显示过程也会有帮助。随着数码相机用户数量的剧增，图像的印刷技术发展得非常快，本书提供了最新的相关信息。

图像测量通常是获取科学数据的主要方法，而这通常又需要定义明确的特征或结构，它们可以是边缘、亮度、色彩、纹理或它们的组合。在确定适当的处理步骤时，对整个场景或对每个特征进行何种测量非常重要。书中的几章详细探讨图像的测量。尽管人类视觉并不是定量的，且很容易被误导，但尺寸、位置和亮度的测量涉及人类能大量理解的主题。形状是更难以理解的一个概念，因此这一版中增加了单独的一章来总结用数字描述形状的各种方式。测量的数据可用于目标的分类或识别。书中给出了几种不同的测量方式，并给出了示例。

类似于食品加工与文字处理，图像处理并不会减少所呈现的数据量，而只是对其进行重新排列，意识到这一点会很有帮助。在这种意义上，有些排列可能会更吸引人，有些排列则可能会表达出更多的含义，但这两个标准可能既不相同，也不会要求相同的方法。

本书介绍了许多图像处理工具，以便基于计算机系统的用户既能理解套装软件中提供的这些方法，又能为特殊应用的需要编写程序。通过从各种显微图像、肉眼可见图像、遥感及天文图像中选取有代表性的图片，比较了用于相似目的的不同算法。要强调的是，图像的尺度对所用来处理或分析的技术影响甚微。具有纳米级分辨率的显微镜与产生光年级图像的望远镜，会生成需要许多相同算法的图像。

贯穿全书的重点仍然是解释和演示方法，而不提供大量的数学公式，以便读者能清晰地理解书中的内容。随着计算机速度与运算能力的提高，搜索有效性的技巧与近似方法的重要性有所降低，因此基于全精度方法的精确实现的例子通常能在桌面系统上实现。所涉及的主题通常与典型工作流程中方法应用的相同顺序呈现。

多年来，在向学生讲授这些内容时，我将掌握这些技术比拟为一名熟练的木匠。不同的木工工具（锯、刨、钻头）的数量相对较少，且尽管存在一些变体，如螺丝起子有平口形的也有十字形的等，但知道如何使用每种类型的工具则与完全理解它们紧密相关。使用一组这样的工具，熟练的木匠可建造出一栋房子、一艘船或一件家具。图像处理工具同样如此，它也分为几种类型，如直方图修改、邻域运算、傅里叶空间处理等，可用来实现许多目的。在五金店中购买工具时，并不提供使用这些工具的技巧。理解怎么使用它们需要实践，从而培养自己预先想象如何使用这些工具的能力。对于图像处理工具而言，同样如此。

在为本书修订这一版时，我也试图对读者与审阅者的评论和要求做出回应。在第二版中添加了关于图像测量及数据解释的章节，在第三版中添加了关于表面图像的一节。第四版在三维结构和特征分类的不同逻辑方法几节中增加了测量的体视学解释。第五版扩充了关于去卷积、扩展动态范围图像和多通道成像以及主成分分析的几节。在第六版中，添加了关于形状含义的一章和多维成像的补充内容。在描述日新月异的图像获取与打印硬件的章节，添加了最新的硬件和软件技术内容。

与前几版一样，我一直拒绝采纳在本书中“添加更多的数学知识”这类建议。对于图像处理、压缩、数学形态学等，已有许多优秀的教材，这些教材提供了严密且丰富的推导，其中的许多也被本书所引用。但本书的要点仍是采用例子的方式来进行教学。很少有人通过公式学习图像处理的原理。就像我们用图像来交流思想与“科学研究”一样，我们中的多数人使用图像来了解许多事情，包括成像本身。通过看到并比较不同运算对图像处理的结果，我们才能发现如何及为何使用这些运算。只有这样，我们才能更好地理解其数学原理。

对于每个涉及成像的人，尤其是从事科学和取证领域的人而言，一个非常关心的实际问题是正确和适当的处理由什么组成，以及不道德的甚至欺骗性的操作由什么组成。简短的答案是，更改一幅图像进而创建就观察者而言的错觉的任何手段都是错误的。该答案所对应的问题并未考虑不同观察者可能希望从图像中看到不同内容这一事实，因此，对某人而言的错觉，对另一个人可能并非如此。

首要原则总是，要存储原始图像及关于其获取的相关数据的永久副本。次要原则是，要小心地记录处理图像的每个步骤，并在发布处理后的图像时报告这些步骤。多数科技出版物及审阅所提交论文的编辑，近年来更意识到进行图像处理的容易性与文档不足的危险性。例如，读者可参阅 M. Rossner 和 K. M. Yamada 的文章（2004；J. Cell Biology）来了解该期刊对于图像伦理的政策及不正确处理的例子。

根据美国最高法院的 Daubert 案件［Daubert 状告 Merrell Dow 制药有限公司（92-102），509 U.S. 579，1993］，出于取证目的，处理人员有责任完整地记录图像处理的每个步骤，以确保这些方法能被法庭接受，这通常意味着这些方法不仅广泛被专业人员接受，而且已被严格测试并输出了想要的结果。在取证环境中，通常需要向非技术性陪审团逐步解释过程。这通常要求证明从图像得到的细节确实在原始状态下是真实存在的，它们只是通过处理在视觉上更加显而易见。

一些过程，例如对特征重新排列或将它们组合到单幅图像中，或调整几幅图像的对比度以使它们看起来更相像，很明显是误导或完全错误的。有些过程，例如使用复制－粘贴来复制图像的一部分，或选择性地擦除图像的某些部分，是完全不正当的。即使选择性地裁剪一幅图像（或选择要记录的范围）也会让人们产生错觉。

基本准则是，对一幅图像添加任何内容都是不可接受的，但若是使得剩余的细节更易于得到，或形象地用于展示与沟通，或是便于测量，抑制或移除某些信息，则可能是可以接受的。当然，所用的过程必须记录下来并报告。这里显示的任何过程可能在特殊情形下是适当的，但它们也可能被误用，在没有对其理解与仔细记录的任何情形下，都不应使用。科学方法的核心是重现性。如果在所应用的处理步骤中提供了足够的信息，且保留了原始的图像数据，那么结果的有效性就可被独立验证。

重要但经常被忽视的一个要点是，需要避免在用户没有意识到的情况下使用程序更改图像。例如，使用 Photoshop 仔细地校正一幅图像的颜色，然后将它放入 PowerPoint 中演示，即使是在同一台计算机屏幕上也会导致不同（在转换过程中使用的复制－粘贴操作同样会丢弃像素并降低分辨率）。此外，在另一台计算机显示器上或使用投影仪时，图像看上去也会有所不同。将一幅图像粘贴到 Word 中会降低图像的分辨率及彩色或灰度动态范围。这可能不会影响打印出的文档，因为打印文档的色域比计算机屏幕的色域要小，但接下来不能再从打印文档中恢复出该图像的原始形式。使用有损压缩方法（如 JPEG）保存一幅图像时，可能会丢弃重要的信息，且这些信息将不可恢复。

我鼓励读者在使用本书的同时，使用真实的图像源和基于计算机的系统来用不同的方法随意进行

实验，以确定最适合自己特定需求的方法。在读者预先不了解相关内容之前，选取图像处理工具来考察图像，与那种使用工具来使得图像更易于被另一位观看者看出与你已发现的相同内容相比而言，是一件更困难的任务。这需要很快的计算速度和界面的交互特性，尤其要求读者成为一名非常善于分析的图像观察者。如果能学会看出计算机所看到的内容并预估各种算法的结果，那么读者将成为一名出色的观察者，并可能获得适合于进一步处理与分析的最好图像。

为使得这种亲身实践的学习更加方便，我与我的儿子Chris Russ合作撰写了本书的配套图书*Introduction to Image Processing and Analysis*，该书介绍了如何编写程序来实现这些算法并创建与 Adobe Photoshop 兼容的实现这些方法的插件。该书的可下载习题解答适用于使用 Macintosh 或 Windows 操作系统的计算机。网站 http://www.DrJohnRuss.com 上还有下载插件和试用程序包的其他链接。

作者简介

在作为科学家和教育家的50年生涯中，John C. Russ一直使用图像处理和分析来作为理解和描述材料结构和功能的主要工具。来自各种设备（包括光学和电子显微镜、X射线和中子断层扫描等）的图像要求计算机处理和测量，以提取重要的数据。Russ的许多研究工作涉及金属和陶瓷的微观结构与表面形貌。他的研究与教学获得了来自政府机关和业界的基金资助。尽管现已退休，但Russ博士目前正协助北卡罗来纳州立大学建立新的实验室和项目，这将是全美第一家提供取证学与工程高级学位的地方。

对算法与设备的熟悉使得Russ博士的技能扩展到了更为宽泛的图像领域——从天文学到生物医学研究，再到食品科学与取证。另外，对于北卡罗来纳州立大学工程学院的学生而言，Russ一直是其研究生学术委员会成员，此外还在材质、浆纸产品、兽医、微生物学、食品科学、考古学等方面与教员们进行协作。向几千名学生讲授原理与方法并向业界客户提供咨询，进一步拓宽了Russ博士的经验及图像处理与分析的应用范围。

退休后，Russ博士曾担任Rank Taylor Hobson这家精密仪器制造商的研究总监。他仍在撰写图书并为许多公司提供咨询（以及刑事与民事案件的专家鉴定），在全球范围内开办图像处理与分析讲习班，并审阅出版物和基金申请书。他活跃于美国显微镜学会、微束分析学会、国际光学工程学会（SPIE）、国际体视学学会，是定量形态学学会的主席团成员、皇家显微学会会士，并受这些或其他组织的邀请进行特邀讲座与培训。2006年11月16日，纽约显微学会因John Russ在显微领域作为计算机辅助显微与图像分析的开发者，而授予他Ernst Abbe纪念奖。

目　　录

第1章　获取图像

1.1　人类依赖图像来获取信息

人类主要是视觉生物。并非所有动物都像人类那样依赖于眼睛来获取周围的绝大多数信息（人类视觉的特性将在第2章中探讨）。日常生活中的这一偏好还扩展到了我们如何去追求更多的技术目标。科学仪器通常会生成图像，而不是生成传递给操作者的声音或气味。在飞向其他行星的空间任务和同样艰难的深海勘探中，相机总是作为其主要部件，这些任务的成功与否，通常由返回图像的质量来判定。这表明，人类以一定的方式拓展了其视觉范围。显微镜和望远镜这样的光学设备可让我们看到比通常小很多或大很多的事物。传感器能检测出电磁波谱可见部分（400～700nm的较窄波长范围）之外的红外线、紫外线、X射线和无线电波，不久的将来甚至会检测出重力波。图1.1给出了一个例子，这是一幅以图像形式表示的射电望远镜数据的图像，其中的颜色表示射电信号中的多普勒频移。这样的设备和表现方式用于进一步扩展成像性能。

非电磁辐射信号也可用来产生图像。使用原子探针"感知"标本表面的新型显微镜可将其数据表示为图像（见图 1.2）。以这种方式收集的数据可表示表面高度与表面形貌，但也可使用探针上的侧向曳力这样的其他信号。低频声波产生声呐图像，而在十亿赫兹频率处，声学显微镜会产生分辨率与光学显微镜类似的图像，但图像的对比度是由声波而非光的衰减与反射的局部变化引起的。图1.3显示了一幅存在近表面缺陷的声学显微图像，图1.4显示了子宫中婴儿的一幅声波图。

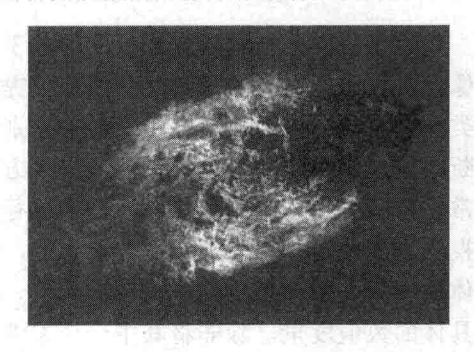


图1.1　射电天文学产生的梅西叶星系图像。这些图像通常以伪彩色显示，以便强调信号强度或多普勒频移的细微变化

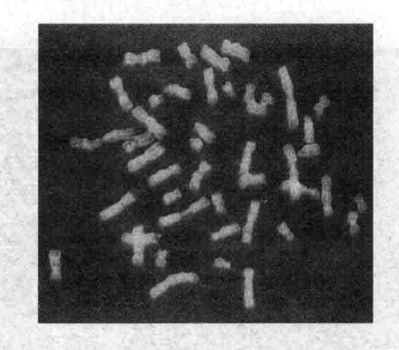
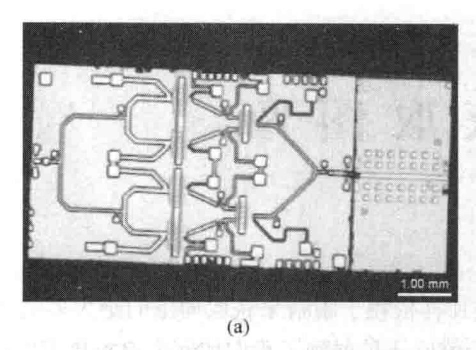
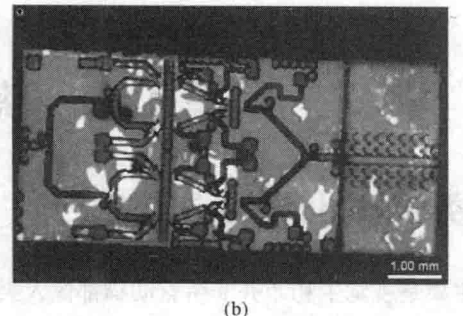


图1.2　人类染色体的原子力显微镜图像

像全息图或电子衍射图样这样的一些图像，所记录的亮度与位置相关，观看者对这类图像并不熟悉。图1.5显示了来自透射式电子显微镜的一幅电子衍射图样的图像，图像中揭示了样本的原子结构（仅适用于进行测量且知道如何解释该数据的人员）。其他类型的数据，包括带有专用符号的气象图、商业利润和开销图，其轴线表示时间、家庭收入、胆固醇级别甚至更多模糊参数的图表，均已成为日常生活中的一部分，如图1.6所示。计算机接口及显示技术的最新进展使图形大量使用，从而进一步利用了人类视觉通路的大带宽。Tufte（1990，1997，2001）特别说明了使用合适图形来传递复杂信息的重要性。



(a)



(b)

图 1.3　砷化镓（GaAs）模具下方焊缝中空隙的声学显微镜图像：(a)模具表面；(b)声学图像，表明空隙表面具有强信号反射（白色区域）



图 1.4　声波图成像的表面重建，所示为子宫内一个 26 周大的胎儿

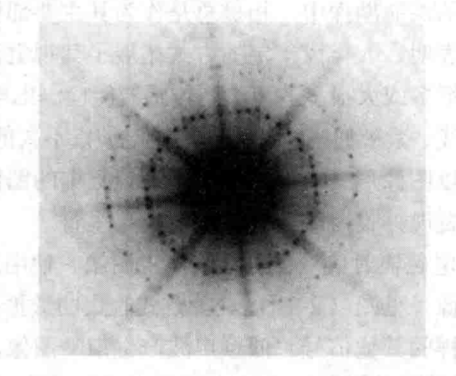


图 1.5　氧化物微晶的会聚束电子衍射（CBED）图样，对其索引和测量，可得到高精度的原子单元尺寸

图 1.6　传递新闻信息的典型图形，包括如股市报告的一维图形和如天气图的二维图形

与用于科学、技术或取证目的的基于计算机的成像设备相比，人类的视觉、从图像中提取的信息的类型以及这样做的原理存在明显的差别。人类在判断图像内的物体和特征的颜色或亮度方面的能力非常差，除非将这些图像排在一起来进行比较。人类视觉本质上是相比较的，而非量化的，它对几个物体的相对大小、角度或位置做出响应，但不能提供具体的数值度量，除非将其中一个参考物体作为量尺。Overington（1976；1992）并不赞同这种广泛接受并已成文的结论，但并未提供令人信服的反驳证据。第 2 章中介绍了人类视觉特性影响其观察内容时的一些结果。

本书并不研究人类视觉通路，但第 2 章中的概述可帮助读者理解人类如何去看事物，从而使得我们成为更好的观察者。在某些情形下，基于计算机的图像处理和分析使用基于人类视觉方法的算法，但也采用看上去与人类视觉无直接对应关系的其他一些方法。特别地，一些图像处理方法基于图像形成和检测过程的物理性质（Sharma，2005）。

书中给出的许多例子和分析涉及源自各种显微镜的图像。通常最直接提供图像的三类成像应用是显微镜、航空（及卫星）成像和工业质量控制。因为在这些情形下，可最大限度地了解和控制成像几何及场景的光照。在更普遍的“真实世界”下，图像内容的分析和解译会变得更为困难。物体离照相机的距离不同，会使得确定其大小变得困难；不同的光照则会改变物体的颜色，甚至会部分地遮掩其他物体。犯罪现场和事故照片需要在苛刻的条件下拍摄，即不是从最优的视点拍摄，而且还伴随着光照的变化，因此对其分析很具有挑战性。

图像处理和测量的基本技术对图像而言是基本相同的，而不管其源自何处及大小如何。源自显微镜甚至天文望远镜的图像，由光子或声波、磁共振或扫描仪生成的图像，有着许多的共性，而处理它们的缺陷的技术、增强和提取细节的技术，以及使用相同算法与技术进行测量等，将在接下来的章节中说明。如后面几章中提及的那样，对于不同几何形状的观察，测量的解译确实要求一些专业知识，但其基本上与放大率无关。

1.2　摄像机

本书第一版于 1990 年出版，当时最常用且可以负担得起的获取图像并进行计算机处理的方法是使用摄像机。将摄像机安装在显微镜或翻拍台上、卫星或空间探测器中，或使用适当的光学设备来观察一个实验，摄像机会将模拟信号发送到一台单独的“帧抓取器”或计算机中的模数转换器（ADC）接口板，然后由该接口板将数值存储在内存中（Inoué，1986；Inoué & Spring，1997）。

原始视频相机的基本形式是摄像机，如图 1.7 所示。其工作原理是，使用聚焦电子束对真空玻璃管内侧的荧光粉涂层进行扫描。光通过前玻璃表面（和一较薄的金属阳极层）进入相机，并在荧光体上产生自由电子。这些操作改变了涂层的局部电导率，因此当电子束扫描时，流向阳极的电流大小会随局部的光强而变化。如图 1.8 所示，这种模拟（连续变化的）电信号被放大，并符合电压和时序标准（欧洲和美国的标准与时序稍有不同，但基本原理相同）。

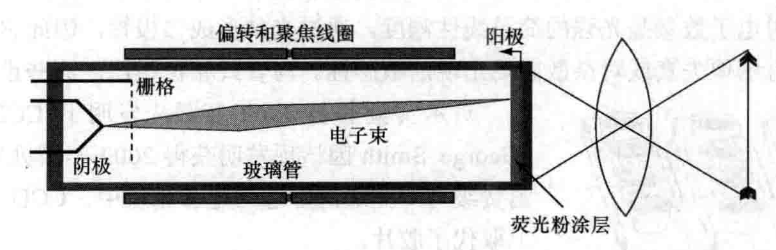


图 1.7　摄像管的功能框图。撞击荧光粉涂层的光改变涂层的局部电阻，进而改变电子束扫描的电流

电压数字化是通过对其采样并生成一个可比较的电压实现的。儿童游戏“猜数字”表明，仅通过 8 次猜测就可以得出 256（最常用的 ADC 类型）个电压值中的某个电压值。第一次猜 128 或该电压范围的一半。如果这个数太大，那么第二次猜测减去 64。后续每一次逼近加上或减去的值等于前一次的一半。在第 8 步做出最终（也是最小）的调整。得到的结果是一个可方便地存储于大多数现代计算机 8 位存储器中的数字。

管式相机既有优点也有缺点。使用电磁场或静电场扫描束会产生一种失真的扫描（枕形或桶形失真，或更复杂的情形），且会受到来自连线或设备的杂散磁场的退化。图 1.9 给出了枕形失真、虚光和失焦的例子。保持图像角落的聚焦要采取特殊的电路，而且角落会因镜头光圈和通光玻璃厚度的下降而变暗（虚光）。密封的真空系统会随时间而变差，且用于吸收气体分子的“吸气剂”在相机垂直方向

使用时，会剥落并掉到荧光体上。相机的响应（电压与亮度的关系）近似于胶片与人眼的对数响应，但它会随亮与暗的场景而变化。从较亮场景和亮点恢复较慢，且在亮光产生的在涂层中横向传播的点处，会出现看上去大于其实际特性的扩散性模糊现象，并在扫描方向出现“彗尾”。

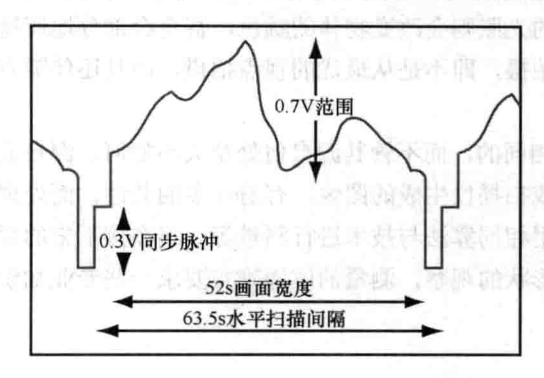


图 1.8　标准的 RS-170 视频信号显示了沿扫描线的亮度变化（范围介于 0V = 黑色与 0.7V = 白色之间）

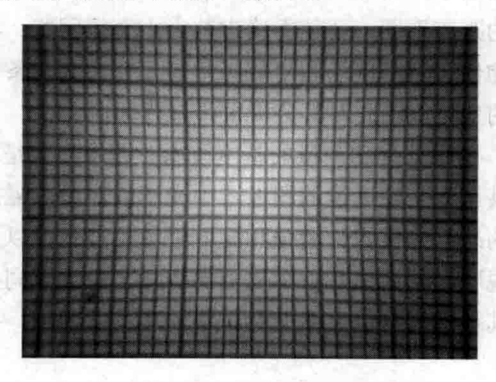


图 1.9　显示枕形失真的图像示例，在边缘和角落出现散焦和虚光

但管式相机也有不少优点。其空间分辨率非常高，仅受荧光体颗粒大小和聚焦束点大小的限制。此外，荧光粉有一种类似于人眼的谱响应，即可看到从红（波长约为 0.7μm）到蓝（波长约为 0.4μm）的颜色。用带有中间阴极层或亮度的特殊涂层对基本相机进行改装，可以在非常暗淡的光线下获取图像（如夜景、荧光显微镜）。

1.3　CCD 相机

管式相机现已基本上被固态芯片相机取代，固态芯片相机最原始且最简单的形式是 CCD（电荷耦合器件）。相机芯片包含一个二极管阵列，其作用是作为光桶。进入半导体的光会将电子从价带抬升到导带，使得电子数量是光强的简单线性测度。通过光刻形成二极管，因此它们有着完美的规则图案，所以没有图像失真或对杂散磁场出现的敏感性。与管式相机相比，这些设备便宜且坚固耐用。贝尔实验室于 1969 年最先发明了 CCD 并申请了专利（George Smith 因该项发明获得 2009 年诺贝尔物理学奖），在消费级与专业级的静态与电影相机中，CCD 现已在相当程度上取代了胶片。

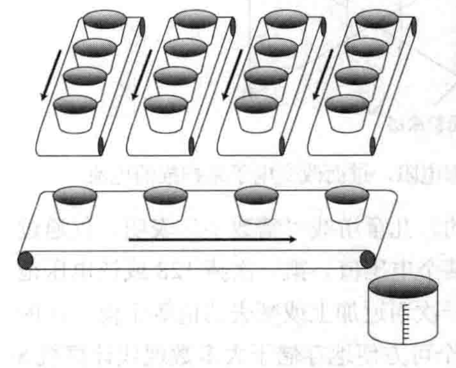


图 1.10　CCD 操作的基本原理可用一组水桶和传送带为例来说明

CCD 的基本操作如图 1.10 所示。每个桶表示相机中的一个“像素”（如下面所解释的那样，在不同的上下文中，该词有不同的含义，因此使用它时要小心一些）。由于在芯片上有几十万个检测器，因此使用线来连接它们进而读取信号是不切实际的。相反，每个桶上由入射光子累积的电子一次一行地被转移到一个读出器中。在一个时钟信号上，每列像素移动一个电荷位置。它将桶的内容放到了读出器中，然后一次一个像素地移动该行，但速度非常快，将电子转储至一个放大器中，放大器产生模拟电压信号，该模拟电压信号可直接送出或测量，进而产生来自数码相机的数值输出。

移动电子最简单的方法如图 1.11 所示。该设备表面上的每组三个电极组成一个像素。通过在其中的两个电极上施加电压，就在作用类似于一个桶的半导体中建立了电场。电子被两侧的高电场困入中心区

域。注意，这并不会减少对输入光子敏感的区域，因为高电场区域中产生的电子会很快跃迁至困住它们的低电场桶中。如图所示，通过在六个步骤或阶段中改变应用到这些区域的电压，电子移动一个像素。首先，一个场区域被降低，电子扩散至更大的体积中。其次，另一端的电场被提升，电子已被移动了 1/3 的像素高度。重复这一过程类似于传送带，并且这就是其称为“电荷耦合器件”的原因。

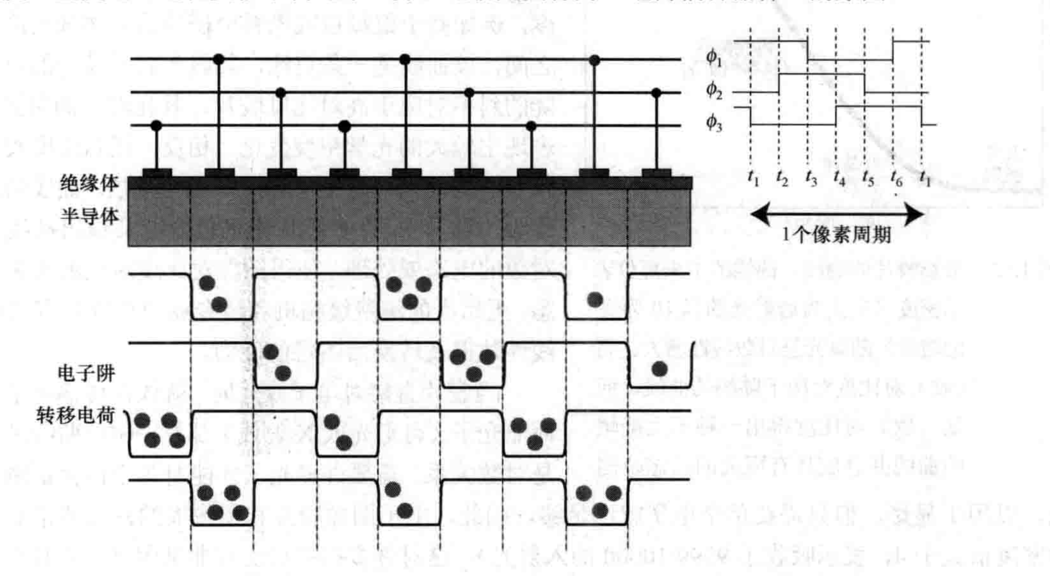


图 1.11　CCD 中三电极组将电子从一个像素移动到另一个像素时的电压变化

芯片相机的一个重要问题是它的谱响应。即使芯片反过来且变细使得光从电极的对面进入，也只有很少的蓝光能穿过半导体进而产生电子。另一方面，红外线很容易穿透，且这些相机对红光和红外（IR）线的敏感性远远超过了人类的视觉，因而通常要求安装遮光滤光片来消除它们（因为红外线不会像可见光那样聚集到同一平面，因此会产生模糊或灰雾图像）。图 1.12 显示了这种谱响应，使用非硅材料，这种谱响应可被进一步“裁剪”和扩宽。使用抗反射涂层时，芯片可实现的较高总效率，主要受“填充因数”（包括保持电气隔离的窄沟之间有源器件的芯片面积比）的限制。此外，芯片相机有一个与入射光强呈线性比例关系的输出，这对于某些测量目的来说非常方便，但这完全不同于人类视觉、摄像机和摄影胶片，它们的输出则与入射光强呈近似对数关系。

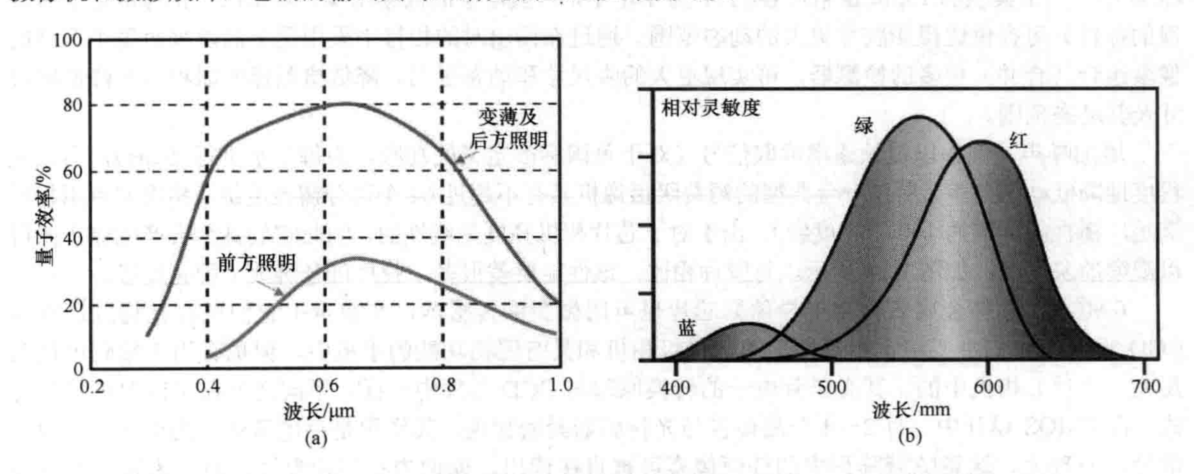


图 1.12　谱响应：(a)硅衬底芯片；(b)人眼中的色彩传感器，它通常识别为红色、绿色和蓝色的敏感性，但涵盖了长、中和短的波长范围

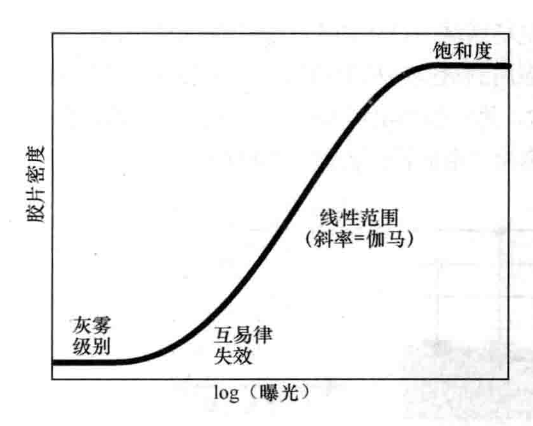


图 1.13 摄影胶片的响应。曲线的中央部位表示密度（定义为透射光的以 10 为底的对数）随曝光呈对数线性增大。高（硬）对比度对应于陡峭的曲线，而低（软）对比度得出一种不太陡峭的曲线并且胶片有更大的动态范围

人类视觉可分辨亮度百分之几的差别。胶片用对曝光的响应来描述，曝光（经化学显影后）产生密度－曝光曲线，如图 1.13 所示。该曲线的底端表示胶片的灰雾级别，即无曝光时的密度。在顶端，胶片饱和到一个最大光学密度，例如基于银颗粒或染料分子的最大物理密度。在两者之间，该曲线是一条直线，其斜率表示胶片的对比度。较陡的斜率对应于高对比度胶片，其在较小的曝光变化时，表现出较大的光学密度变化。相反，低对比度胶片有一个较宽的范围来记录更大亮度范围的场景。曲线的斜率通常称为“伽马”。许多芯片相机包含将其输出从线性更改为对数的电路或处理，使得图像对比度对观察者来说更加熟悉。更昂贵的消费级相机和最专业的相机具有读取“原始”线性数据及转换后图像的能力。

当胶片直接对电子曝光时，就像在传输电子显微照片而非光子（可见光或 X 射线）那样，响应曲线是线性而不是对数关系。需要许多光子才能对各个卤化银颗粒完全曝光，以用于显影，但只需要单个电子就已足够。因此，电子图像胶片和图像板的密度通常非常高（光学密度值大于 4，表示吸收了 9999/10000 的入射光），这对许多扫描仪而言非常困难，并且需要用大于 8 位来进行记录。

相机芯片的发展趋势是体积更小而像素数或二极管数越来越多。一些科学相机，如哈勃望远镜上使用的相机，占据了整个晶圆。但对于消费级设备，每个芯片在整个（对角）尺寸上为 1/3 英寸、1/4 英寸甚至 1/5 英寸，在单个晶圆上放置许多器件可得到更大的经济产量，还需要更小更便宜的镜头。在减小了的芯片面积中放入更多的像素（对于更高的空间分辨率，见下面的讨论），会使得单个检测器变小，但其间的沟必须维持在阻止电子横向扩散的相同尺寸。结果是明显地降低了整个效率。有些设备会将较小的镜头放到二极管上，以便捕获那些另外落入沟的光，但这会增大成本，还会增加和二极管本身一样的不均匀性（整个芯片上通常在 1%以内）。

令检测器更小的另一个且更为重要的效果是降低了它们的电子容量，称为阱容量。在科学级别的 CCD 中，一个典型的 15μm 像素可容纳 50 万个电子，一些电子的低读出噪声（就像在特殊情形下可实现的那样）可提供比摄影胶片更大的动态范围。通过在读出时的相移中采用更多的阶来对每个存储的像素组合（合并）更多的检测器，可实现更大的阱尺寸和动态范围。降低检测器的面积也会降低阱尺寸及其动态范围。

增加噪声，例如以视频速率读取信号（对于美国标准定义的视频，为每个水平行 52μs），会很大程度地降低动态范围，所以一台典型的消费级摄像机具有不超过 64 个可分辨亮度级（昂贵的专用相机满足广播视频标准的 100 个亮度级）。由于对于芯片相机亮度是线性的，因此它们甚至会产生更少的可以观察的灰度级，如图 1.14 所示。与胶片相比，该性能要差很多，胶片可分辨数千种亮度级。

CMOS（互补金属氧化物半导体）芯片也可用做图像传感器，在数量上它们现在已超过最初的 CCD 设备。它们主要用于相对便宜的消费级相机和具有照相功能的手机中，但也有用于数码单镜头反光（单反）相机中的。其光子到电子的转换原理与 CCD 芯片中一样，不同之处在于读取信号的方式。在 CMOS 设计中，有 2～4 个晶体管与光传感器紧密相连，其作用是将电荷转换为电压，并放大信号。原理上，这意味着阵列中的任何像素可被直接读出，就像内存芯片那样，对像素按行和列寻址（见图 1.15）。这不同于 CCD 所用的方法，CCD 所用的方法是将电荷“扫到”阵列的一角，并按固定的顺序读取所有像素。

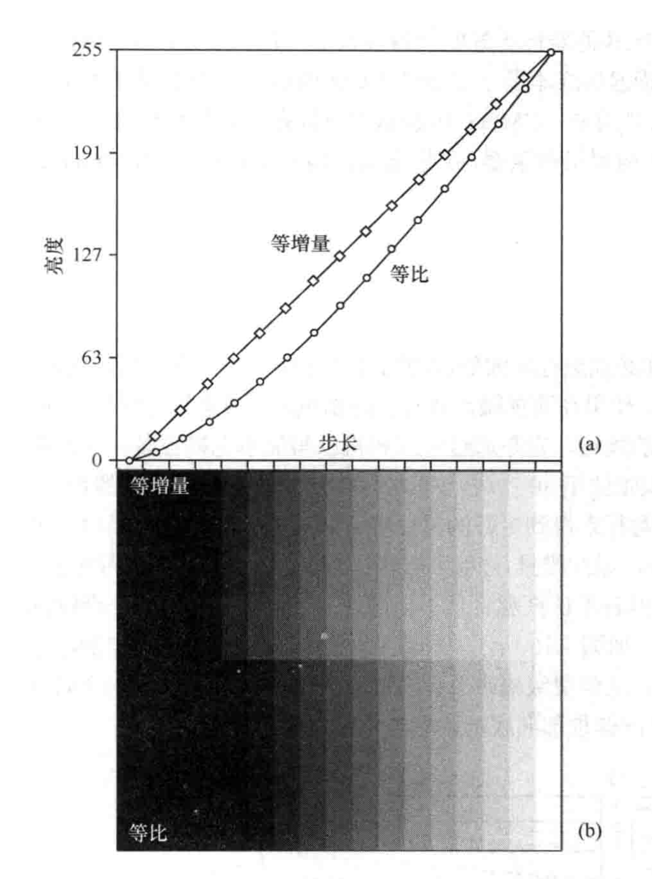


图 1.14　线性（等步长）检测器和对数（等比）检测器的灰度级比较：(a)亮度图；(b)图(a)中数值的显示

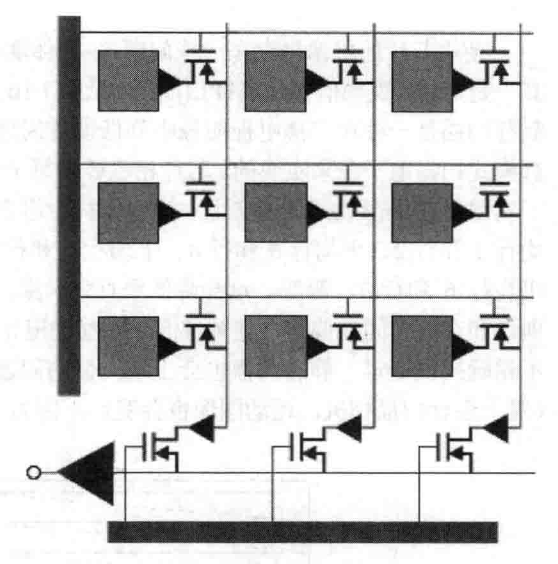


图 1.15　典型 CMOS 检测器的示意图。每个有源光传感器（绿色）有与寻址和输出线相连的额外晶体管

这些控制晶体管所占用的空间减少了“填充因数”或对光敏感的芯片的有效面积，但通过将镜头放到每个检测器的上面来收集来自死区的光并将其导向到活动传感器，可对其进行补偿。这些镜头和为每个像素使用的单个放大器，通常会使得 CMOS 检测器中的传感器不如 CCD 阵列中的传感器均匀，因此会产生需要使用软件进行补偿的固定模式（要求使用不均匀的光照来记录一幅图像）。除了固定模式噪声外，CMOS 检测器通常会有更多的随机噪声叠加到图像信号上，这是由单独的放大器、额外的连线、与之相关的电容和热噪声以及更大的暗电流引起的。非常小的活动区域（至少在袖珍相机和电话使用的小芯片中，尤其是当像素上升到几百万个时）有较小的阱容量，进而导致图像的有限动态范围。图像通常以每通道 8 位进行存储，因为其内存是以传统方式组织的，但通常没有如此丰富的实际亮度分辨率。

由于制造出了具有较大检测器的较大面积的 CMOS 芯片，因而有更大的阱容量和动态范围。当用于更昂贵的相机中时，CMOS 这种设计的优点之一是，存取像素的电路可沿阵列的两个相邻边放置（分别对行和列寻址）。这就使得在另两边可以对芯片进行小心的修剪，将 4 个芯片排列在一起可生成具有更高像素的较大传感器。这种方法与较大的传感器组合使用来达到更大的敏感度和动态范围，已使得一些制造商愿意对数码单反相机采用同传统底片一样大的 CMOS 检测器。

CMOS 传感器的优点主要在于三个方面：消耗更少的能量，因此电池供电时间更长；放大器和数字化电路可放在同一芯片上，从而减小尺寸并增加耐用性；晶圆的生产方法基本上与其他标准硅器件（如存储器和处理器）的生产方法相同，而 CCD 晶圆需要独特的处理。后一个优点在某些程度上被如

下事实抵消，即高质量的 CMOS 传感器若要做到比其他器件更好的一致性及更少的由噪声产生的缺陷，需要一定程度的定制加工。尽管制造 CMOS 传感器的成本要小于制造 CCD 的成本，但其设计成本非常高。当然，对于批量制作的器件，这是一个次要因素。CMOS 传感器的整体发展趋势是继续提升质量与性能，但同时 CCD 传感器的优点对多数技术应用仍很重要，可根据具体情况进行权衡（Nakamura，2006；Holst & Lomheim，2007）。

1.4 相机缺陷与限制

使用芯片的摄像机存在一些问题，主要体现在必须进行后续处理的图像中存在一些特殊类型的缺陷。其一是，许多视频信号是隔行扫描的（见图 1.16）。使用高清视频及数码静态相机时，图像是逐行扫描的。隔行扫描是一种在广播电视图像中降低视觉闪烁的技巧，它是通过管式相机以与显示电视设备相同的隔行模式扫描电子束来实现的。芯片相机要求每个偶帧的第 30 次时，读取阵列两次，一次用于收集偶数行，一次用于收集奇数行。事实上，许多相机会组合两行来得到更好的灵敏度，即在一个隔行扫描场中，平均行 1 和行 2、平均行 3 和行 4、平均行 5 和行 6，以此类推，然后平均行 2 和行 3、平均行 4 和行 5、平均行 6 和行 7，等等。这会降低垂直分辨率，但若不认真观看是不会发现的。运动会导致一个全帧的偶场和奇场之间的偏移，使图像产生明显的退化，如图 1.16 所示。如果水平回扫信号对于要锁定的电子不精确或很困难，静态图像也会出现类似的问题；从消费级磁带录像机回放信号时，这尤其是一个问题（对于逐行扫描相机，运动图像也会变形，因为从图像顶部向底部读取需要时间）。

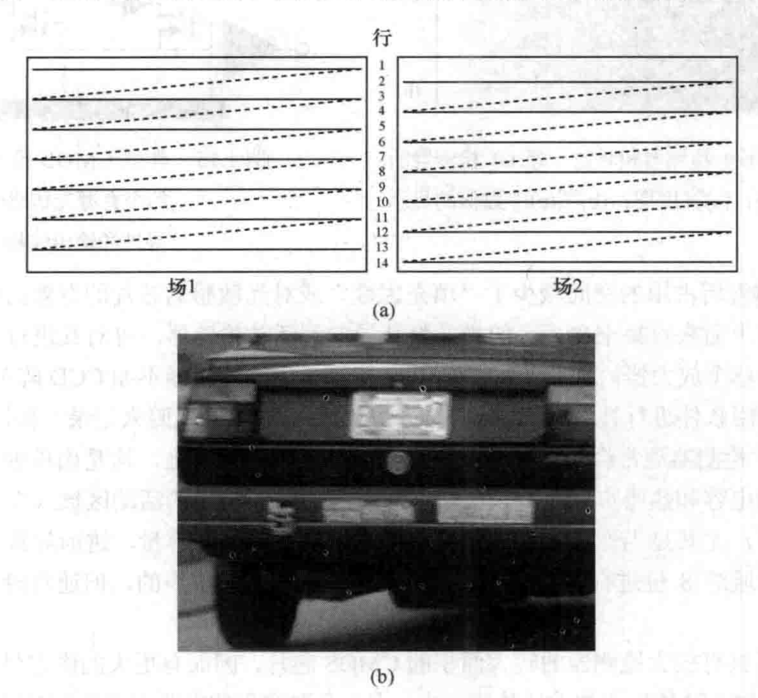


图 1.16 (a)隔行扫描在一个 1/60 秒场中涵盖偶数行，而在第二个场中涵盖奇数行；(b)出现运动（无论是场景移动还是相机移动）时，图像中会出现偏移

在传送与读出过程中，除非相机采用机械或电子快门，否则光子会继续在芯片上产生电子。这会产生较大的背景信号，从而进一步降低动态范围并可能产生模糊。电子快门通常一次完成一行，因此活动图像就变形了。一些设计通过使像素数加倍来避免快门问题，即一半像素挡住入射光。单次传输将电子从活动检测器移到隐藏检测器，从后者读出它们。当然，这会使得芯片上设备的有效面积（填

充因数）的灵敏度降低 50%。

水平行高速读出会产生信号的水平模糊，又一次降低图像分辨率。部分原因在于没有足够的时间让电子沿移动场来扩散，在于从陷阱（硅晶格中的杂质）中恢复电子所需要的时间，而部分原因在于放大器频率响应不够充分，这是对放大器噪声的折中。即使单个电子传输非常有效，在大多数情况下超过 99.999%，但在被收集并放大之前，通过许多这样的传输的结果是增加了噪声。这从芯片的一侧变至另一侧、从顶部变到底部，并且如果没有许多其他细节或运动来模糊它，则可在图像中清楚地检测到它们。

从一个检测器到另一个检测器的许多电子传输发生在芯片读取期间，这会在信号中产生噪声。电荷产生与收集中的纯统计变化相对影响较小。小电荷至电压的转换及其后续放大是大多数系统中的主要噪声源。减慢传输过程可降低读取和放大器噪声，从而在移动过程中丢失更少的电子，且放大器时间常数可整合出更多的噪声，生成更清晰的信号。将芯片温度降至–40℃ 也可降低来自这些因素和暗电流或热电子的噪声。当然，减慢读出速度及降温仅用于非视频应用。数码静物相机使用相同的芯片技术（但使用更多数量的检测器）作为固态视频相机；而且因为读出速度较慢，会产生更高质量的图像。Janesick（2001）探讨了各种噪声源，以及如何对在天文成像（在其中几乎完全替代了胶片）及空间探测中使用的科学 CCD 的噪声进行控制。

1.5 彩色相机

彩色相机可以用三种主要方法设计，如图 1.17、图 1.18 和图 1.19 所示。对于静态图像（其包含许多科学应用，如显微镜，但不包含如视频这样的实时应用），单个检测器阵列可分别用来通过红色、绿色和蓝色滤镜获取三个连续曝光（见图 1.17），然后将曝光组合到一起进行查看。这种方法的优点是低成本，对不同的颜色波段可使用不同的曝光时间，进而补偿硅片对短波长（蓝）光的灵敏度的不足。

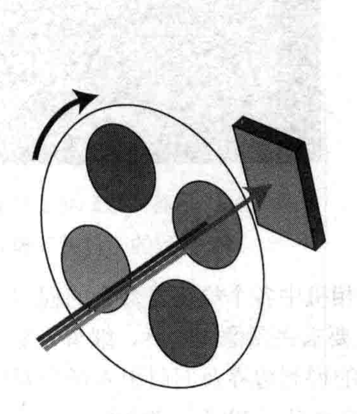


图 1.17 带有红色、绿色和蓝色滤镜的色轮相机的示意图。第四个滤镜的位置是空的，使相机可以作为一个对昏暗图像（如荧光显微镜）有更大灵敏度的单色检测器使用

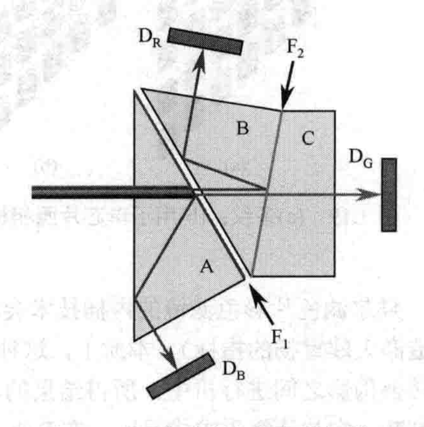


图 1.18 三芯片彩色相机的棱镜和双色向滤光镜示意图

许多高端消费级和多数专业级视频相机使用三个传感器（见图 1.18）。棱镜阵列将入射光分为红、绿和蓝分量，它们由三个不同的传感器记录，其输出以电子方式被组合在一起，产生一幅标准的视频图像。因为需要三块芯片，因此这种方法成本高，但对于视频应用，它们不需要特别高的分辨率（即使高清视频相机，其像素数也比数码静物相机的像素少）。为保持所有东西对齐的光学和硬件也会增加成本，棱镜光学器件的深度使得使用短焦距（广角）镜头不实用。在数码静物相机中，这种设计很少使用。

视频图像通常数字化为640×480的阵列来存储像素（曾是个人计算机的标准VGA显示尺寸），但这不是图像的实际分辨率。广播带宽限制了高频并消除了亮度和颜色的快速变化。标清视频图像水平方向的亮度（光强）信号的实际分辨率不超过330，是色彩（色度）信息的实际分辨率的一半。色彩信息会在分辨率上刻意降低，因为人类视觉对超过边界线的彩色的模糊不是非常敏感。

当然，视频信号会被较差的设备进一步降质。在消费级磁带机上记录视频会使分辨率再降低50%或更多，尤其是在磁头较脏或磁带已被使用多次的情形下（监控视频的取证调查中，一个很遗憾的常见问题是，磁带已被当地警察反复播放多遍，当专业人员拿到磁带时，氧化物涂层和信息已被毁坏甚至删除）。尽管HDTV（高清电视）已在某些程度上有所改善，但视频图像的分辨率仍然不高。因此，除非有特殊要求来获取实时图像（即25～30帧每秒），视频技术通常不是科学成像中的最佳选择。数码静物相机已在很大程度上替代了它们，因为数码静物相机可产生更高分辨率且有着更大动态范围的图像。

多数数码相机使用单像素阵列［通常有很高的像素（检测器）数］和一个允许红光、绿光、蓝光进入特定检测器的彩色滤镜。可以使用不同的模式（见图1.19），拜尔滤色片最为常见［由柯达公司的研究人员Bryce Bayer发明，这一发明基于1976年授权的美国专利（专利号3971065）“Color Imaging Array”］。注意，相对于红色或蓝色，它为绿色安排了两倍数量的检测器，模仿人眼对绿色的更大敏感度。当然，单芯片相机的问题是，每种颜色通道的图像分辨率会降低。例如，在某些位置的红色亮度必须由邻近的传感器内插得到。还需要设计滤镜来在每个通道中提供相同的亮度敏感性。如果这样做得不好，会在均匀灰度测试卡的图像中出现人字形的图案（通常称为“拉链”），且在图片中沿对比度边缘会出现彩色条纹，如图1.20所示。

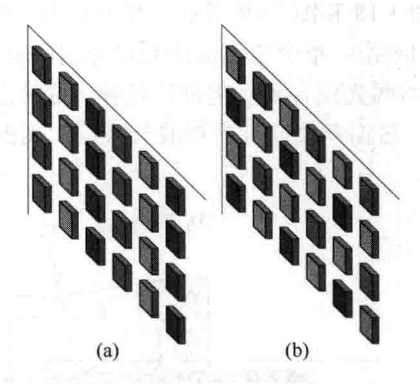


图1.19　(a)条纹；(b)用于单芯片照相机的拜尔滤色片

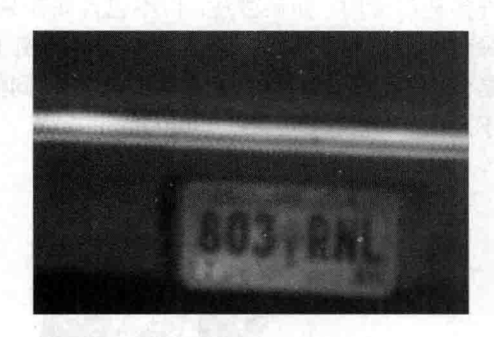


图1.20　单芯片数码相机中较差内插导致的“拉链”模式的例子

拜尔滤色片彩色滤镜的内插技术会将图像分辨率降至相机中各个检测器数量的程度（后者通常是制造商大肆宣扬的指标）。本质上，这种“去马赛克”处理要求在图像清晰度、细节、噪声、处理时间和转换伪影之间进行折中。所得结果的质量由其保持亮度的锐利边界及同时引入最少彩色伪影的能力来判断，它与计算需求成反比。在Ramanath（2000）、Shao et al.（2005）、Tamburino et al.（2010）和Guanara et al.（2010）中，可找到几种专利方法的比较。特殊彩色滤镜阵列排列的组合及相机的内插固件会在图像上留有签名，在某些情形下，这些签名可用来识别用于拍摄场景的相机的型号，甚至识别来自某台相机的特定固定模式噪声，并检测以后对图像做出的变化（Bayram et al.，2006；Swaminathan et al.，2007；Farid，2008）。

模式噪声并非仅出现在带有彩色滤镜阵列的单芯片相机上。三芯片相机也会出现问题，因为所有的芯片在来自各个晶体管的输出方面都有一些微小的变化。在三芯片系统中，这些变化会产生红、绿、蓝输出的不同变化，进而增强图像的色彩变化。

由Foveon公司开发并用于某些相机中的另一种彩色相机设计方法，使用了CMOS技术，它在每

个像素位置生成三个堆叠在一起的晶体管。蓝光在硅中的穿透距离最短，因此被最上面的晶体管检测到。绿光穿透第二个晶体管，红光穿透最底部的晶体管。输出信号组合在一起就产生了色彩信息。这种方法由于采用了内插技术而不会降低空间分辨率，但存在一致性或精确的色彩保真问题。

1.6 相机分辨率

即使是在相机机身内或在同一芯片内进行了数字化，来自硅检测器的信号是模拟信号，因此内插在放大器阶段进行。多数情形下，具有单芯片相机和滤镜排列的实际图像分辨率只是广告号称分辨率的 1/2 或 2/3，其原因就在于这种内插。有些相机会以比芯片分辨率高得多的像素来存储图像。这样的内插和空放大对图像并不能贡献额外的信息。

基于实际分辨率而非声称的记录像素数目来比较相机较为困难。因此，理解“像素”一词的多重意义很重要。在某些情形下，像素是指相机中光检测器的数量（不考虑任何彩色滤镜，有时包括在边缘附近对实际图像无作用但用于测量暗电流的检测器）。在有些情形下，它描述的是存储在计算机中的亮度或彩色值，尽管这可能表示空放大。在另外一些情形下，它用于描述计算机显示器上显示的颜色点数，即使图像是以压缩方式或放大方式显示的。在考虑实际图像分辨率时，区别这些不同的含义来谈论分辨率才会有意义。这指的是整幅图像能彼此区分的离散点数，有时按照可分辨的线对数目来指定。线对数目为分辨率元素的 1/3～1/2，因为一条线至少需要一个元素，线之间的间隔也至少需要一个元素。它取决于线和间隔之间亮度对比度的数量以及出现在图像中的噪声量（随机变化）。

对于一些数码静物相机而言，情形会更复杂一些，因为静物相机会移动检测器阵列来获取图像的多个样本。最常见的方法是使用一种压电设备将阵列在水平和垂直方向上偏移半个像素间距，由于得到的数据来自于原像素位置间隙，因此得到的 4 幅图像可以组合起来使得图像的分辨率增加大约一倍。对于使用彩色滤镜的阵列，额外的移动会产生分辨率接近相应像素间距的彩色图像。一些专用相机会多次将整个传感器阵列移动到胶片平面的不同位置来收集拼贴，随后组合到与检测器阵列大小相同的一幅图像中。当然，这些方法需要多次曝光，也就意味着获取图像需要更多的时间。

除二维检测器阵列外，也可以使用线性阵列（或有时使用三个阵列，分别用于红色、绿色和蓝色滤镜），它扫描整个图像平面来获取数据。这种方法通常用在台式扫描仪中（对许多应用而言这是一种完美获取图像的设备）。它也用在专用相机中，且一些光学显微镜也通过移动光学器件下的镜台和标本来以更高的空间分辨率得到整个 1×3 英寸的幻灯片，从而实现同样的目的。产生的图像文件非常大；因此要求使用专用软件来有效地存取所存储的阵列（Bacus & Bacus，2000，2002）并在用户改变位置和放大倍数时，交互地传送图像数据的所选部分。网络存取这样大的存储图像会面临带宽问题，但便于协作与教学。

不管是使用单片还是使用三片相机，蓝色通道中的噪声通常最大，因为底部的芯片对蓝光敏感且后续需要更大的放大。在许多情形下，使用均值或中值滤波器（第 4 章中将对其进行探讨）以降低图像噪声的处理软件可单独用于每个彩色通道，根据实际的噪声内容使用不同的参数，从而可最佳地改善图像外观。

与摄像机使用相同芯片技术的数码相机可产生更高质量的图像。部分原因在于更长的曝光时间，因为更长的曝光时间会收到更多的电子，进而因统计与放大来降低噪声。此外，从芯片中较慢读出数据（通常需要 1 秒钟或更长时间，而不是 1/60 秒）会降低读出噪声。数码静物相机逐行扫描而非隔行扫描来读出数据。使芯片降温并使用放大电路来降低暗电流，可使用整合（几十秒的长时间曝光，或对某些天文应用采用几分钟的时间曝光）来获得更好的效果，因为一些芯片设计具有较高的动态范围（大的阱尺寸和数字化仪中的大量位数）。此外，使用物理而非电子快门的能力可简化芯片电路并增大填充因数。视频相机中的像素数不需要大于视频信号的分辨率（如前所述，视频信号的分辨率非常低）。

在数码静物相机中，非常大的像素数可提供相当高的分辨率，在有些情形下是对胶片的挑战。

高端消费级和专业科学级相机之间也存在有趣的交叉。除了作为显微镜或其他单独光学器件的专业相机外，制造商正在生产有着足够分辨率（写作本书时为 1500～2000 万像素[①]）的消费级单反相机，因此在技术领域它们也会逐渐变为现实，且简单的光学附件使得它们易于连接到显微镜或其他设备上（当然，相机也可移除以用于其他目的）。相机可直接连接到计算机上，但在许多情况下，先将图像记录到存储芯片上，之后下载到计算机中可能更为实际。有着大和高分辨率检测器阵列、可更换镜头的专业数码相机，可提供与传统的 35mm 相机及较大的胶片型相机类似的性能。随着新的进展频频出现，每家相机制造商已意识到现在是时候从胶片过渡到数字记录了，并且目前有相当多的相机可供选择。

带有大量传感器（高像素数）或较大独立传感器（大阱尺寸及由此得到的高动态范围）的相机的优点看起来很明显且值得期待。对于某些应用而言，高像素数并不重要。在高级光学放大中，重要的限制是光学分辨率。在我的台式显微镜的典型设置中，使用 10 倍（低放大率）目镜，由传输光学元件投射到芯片上的图像大约覆盖标本的 1600μm 宽度。使用 100 倍（高放大率）目镜，则变成 160μm 宽度。对于带有 3600×2400 个传感器（小于 10 兆像素）的相机，低放大率图像约以 1 个像素每微米进行记录，足以满足光学分辨率的要求。高放大率图像以 90 个像素每微米进行记录。由于最优条件下的 100 倍显微镜的光学分辨率约为 0.5μm，因此这会产生大量且不必要的过采样。在低放大率条件下，或是查看大场景（如航空和卫星照片）的细节时，高像素数才有意义。当分辨率的限制依赖于光学元件时，高像素数就可能没有意义了。

1.7 对焦

不管采用何种相机来获取图像，正确地对光学器件对焦来获取图像的细节都很重要。通常，人眼用于人工地执行这一任务。有些情形下，如病理切片的自动显微镜检查或车辆的监视跟踪，都需要自动对焦。这就将计算机处理引入到了图像获取的最初步骤中。有时，为了速度，这一处理由相机附带的专用硬件电路执行。但在许多情况下，算法与计算机中的算法相同（将在第 5 章中描述），而对焦在软件中通过在一个设定范围内对光学器件步进调整，并选择给出最佳图片的设置来完成。

自动对焦可使用几种不同的方法。用于宏观场景的相机可能采用远距离测量技术的方法，如使用高频声波或红外线来求出至目标的距离，以便调整镜头位置。在微观应用中，这是不切实际的，必须使用图像本身中所获取的焦距调整。有多种算法用来检测图像清晰度的质量，在有较好的对比度和细节时，所有这些算法对多数图像都是成功的。每种方法选取高通滤波器的某些实现，该滤波器能以不同方式（如硬件或软件）来实现，但必须考虑图像中的高频噪声及光学器件的光学传输函数的影响（Green et al.，1985；Firestone et al.，1991；Boddeke et al.，1994；Sun et al.，2004；Buena-Ibarra，2005；Bueno et al.，2005；Brazdilova & Kozubek，2009；Shim et al.，2010）。

1.8 电子和带宽限制

固态芯片或管式视频相机都产生对应于图像中不同点的亮度的模拟电压信号。在标清 RS-170 信号转换中，电压从最小亮度至最大亮度的变化超过 0.7V，如图 1.8 中的上图所示。扫描名义上是 525 行每全帧，用两个隔行 1/60 秒场组合来建立整个图像。仅约 480 个扫描行能用，剩余行在垂直回扫时丢失。在典型的广播电视图像中，由于过扫描，丢失的行比这还要多，在实际的可视区域中仅保留约 400 行。

① 本译稿完成时，市面上已有大量 2000 万像素以上的消费级单反相机，如尼康 D3X、佳能 5D Mark III 等。——译者注

每次扫描行的持续时间为 62.5μs，其中部分时间用于水平回扫。这为图像数据留下了 52μs，它必须再细分为分辨像素的水平间隔。对于 PAL 电视，基于 1/25 秒帧时间和更多的扫描行，这些值稍有不同，得到的分辨率要稍高一些。

在美国，直到最近，仍只提供 4MHz 带宽给广播电视台的信号，该信号必须携带色彩和声音信息及亮度信号。这一较窄的带宽限制了单独电压值的数量，即如上所述，沿每个扫描行可分辨最大的值为 330，而且如果该信号被电子或使用标准磁带录像机的录制降质，则该值会降低。消费级质量的磁带录像机会明显降低有效分辨率；在“定格”回放时，它们仅显示两个隔行扫描场之一，所以在垂直方向上仅有约 220 行被还原。使用这样的装备作为图像分析系统的一部分时，若根据分辨率（实际上是采样的像素数）来选择相机或数字化仪，就不恰当了。

标清电视中的隔行扫描与非隔行或“逐行”扫描完全不同。因为不存在行行对齐或偏移问题，后者可提供更好的质量。大多数高清电视（HDTV）使用逐行扫描方式。该格式要求更高的帧重复率来欺骗人的眼睛，以便人眼看到没有闪烁的连续画面，但它还有许多其他的优点，包括更简单的逻辑电路来从相机中读取数据（可能直接集成到芯片上），由于相邻行间的冗余使得有更多机会进行数据压缩，以及更简单的显示和存储设备。实际上，所有的科学成像系统如数码相机、直接扫描显微镜［扫描电子显微镜（SEM）、扫描式隧道显微镜（STM）、原子力显微镜（AFM）等］、平面扫描仪、胶片或幻灯数字化仪和类似的设备都使用逐行扫描。

除使用逐行扫描外，HDTV 模式与传统电视还有许多不同。其像素密度更高，具有更宽的宽高比 16:9（而不是 NTSC 电视中使用的 4:3），且像素是方形的。典型的 HDTV 模式以每秒 60 全扫描的速率展示 1920×1080 像素的图像，总数据传输率超过 2Gbps，是模拟广播电视的一数百倍。高数据传输率直接导致了使用数据压缩技术（这将在第 3 章中介绍）以及使用电缆或光纤而不是广播信道的数字传输技术的应用。广播、卫星和电缆传输的 HDTV 信号都会涉及压缩，这通常会导致图像质量的明显下降。

抛开对消费级电视的影响不谈，HDTV 硬件的发展可能会对计算机成像技术产生附带影响，如带逐行扫描输出、高带宽录制设备以及超级 CRT 或 LCD 显示的高像素密度相机。例如，在相机内进行模数转换，其中对红色、绿色和蓝色的每一个至少用 10 位表示，设计用于 HDTV 应用的彩色相机输出数字而非模拟信息。

电缆、连接器或不正确的终端阻抗这些简单的元件，都会严重降低系统的性能（甚至最好的系统也不能幸免）。使用标准相机的另一个实际需要注意的问题是，避免自动增益或亮度补偿电路，它们会更改对应于亮或暗区域的图像的对比度或线性，而这些区域甚至不在图像的数字化部分，进而使得图像间的对比变得困难，并增大微弱信号的增益和噪声。

图 1.21 显示了一张显微照片及其亮度直方图。直方图是图像分析中的一种重要工具，它画出的是像素数量与像素亮度值的关系图。后续章节中会广泛使用它。所示直方图在可用的 256 个亮度级间分

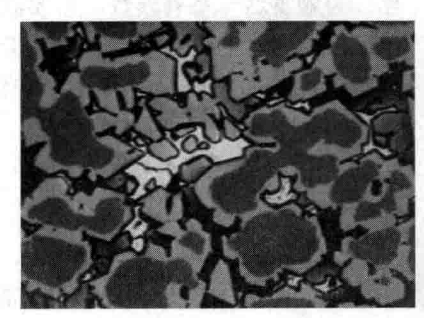
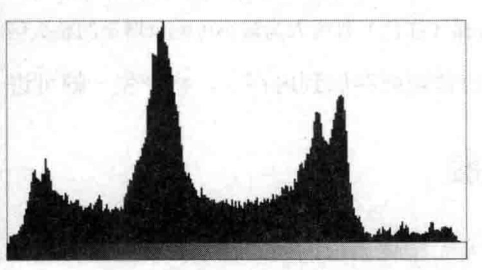



图 1.21　数字化一个显微金相得到的灰度图像及其亮度直方图，直方图画出了像素数量与亮度值的关系图

布良好，峰值对应于金属样本中的每个结构。如果亮灯照射到固态相机中非我们所关心区域的检测器部分（由光学元件中的内部反射引起），那么相机中的自动增益电路可能会改变亮度与电压的关系，以使图像变化。在使用白的或暗的模板环绕置于翻拍台上的一台相机下方的图像时，也会出现相同的效果。结构与亮度的关系是变化的，这就使得后续的分析更为困难。

下面探讨色彩校正与校准，但由于通过每个滤镜的光的大量的波长、光照颜色的变化（甚至白炽灯泡上轻微的电压变化），以及色彩信息编码的方式，由视频相机获得绝对的色彩信息是不可能的。这就要求进行校准来匹配颜色，以便人们对色彩的感觉保持一致，这将在第 4 章中讨论。

所使用光源的色温对匹配图像中的颜色来说至关重要。图 1.22 显示了一幅已滤掉了日光方式记录的图像，日光的有效色温（详见第 3 章）大约为 5000K，方法是使用一张白卡，并提前曝光来让相机进行色彩平衡调整。使用不同的色温打开原始图像文件会在颜色的视觉感知方面产生相当大的变化。

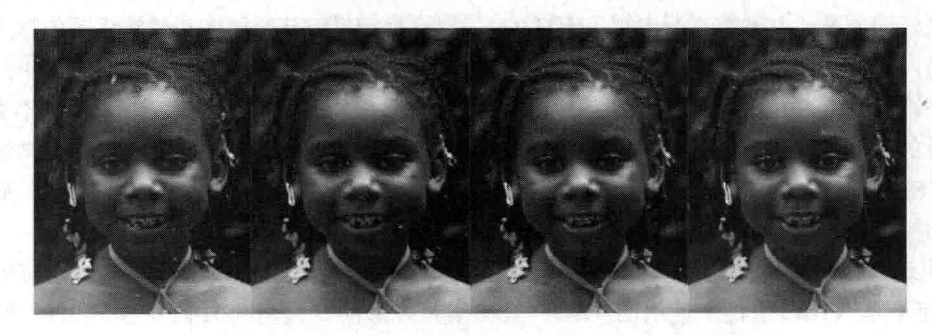


图 1.22　滤掉日光后以有效色温 5000K 拍摄的图像，这些图像分别以 3500K、4500K、5500K 和 6500K 的色温打开

来自检测器的模拟电压信号的数字化可在相机内部完成，也可在单独的外部电路中完成（如置于计算机内部的“帧抓取器”板）。模拟信号通常使用“快闪”ADC（模数转换器）进行数字化。这是一块芯片，它使用逐次逼近技术（见上面的描述）来快速采样并测量电压。对于视频速率成像，这须在 100ns 内完成，产生 0～255 之间的一个值来表示亮度。慢速读出允许超过 8 位的转换，且许多数码静物相机有 12 位甚至 14 位 ADC（尽管检测器内部的动态范围和噪声电平可能没有这样好）。亮度数存储在内存中，并做另一次读取，以便沿每个扫描行获得一组亮度值。图 1.23 说明了在时间和数值上等步长对信号进行数字化。为触发这一系列读数，需要额外的电路，以便沿后续行的位置是一致的。沿每个扫描行数字化几百或上千个点，对每一行重复该过程，并将所得的值传输或存储到内存中，将产生一幅可进一步处理或分析的数字化图像。

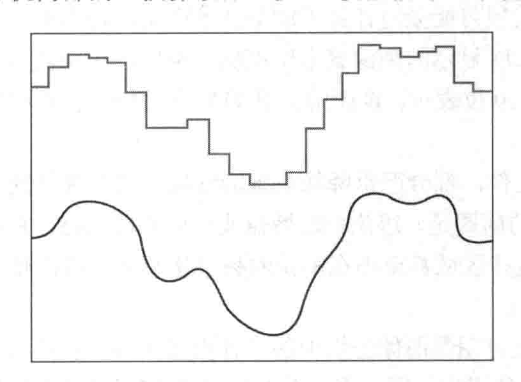


图 1.23　在图像中沿一行对模拟电压信号（蓝色）数字化会产生一系列数值，这些数值对应于时间上的一组阶梯（红色）且舍入为最小可测量增量的整数倍

1.9　像素

我们非常希望像素值的间距在水平和垂直方向上是相同的（即方形像素），因为这会简化许多处理与测量操作。将像素排列成六边形网格在理论上具有优势，但由于所有获取硬件实际的工作方式，并

为简化计算机内存中像素的寻址，这种网格将不会采用。

要使用模拟视频相机实现方形像素，要求有一个已精确校正的时钟来控制获取过程。由于标清视频图像不是方形的，而是有着 4:3 的宽高比，因此数字化后的图像可能只会显示整个视场的一部分。数字化器（帧抓取器）首先设计用来记录 512×512 阵列值，因为 2 的整数幂大小会简化设计与内存寻址。后面获取了一个 640×480 的阵列，它与标准 VGA 显示器的图像比例及尺寸匹配，同时可使像素为方形。由于时钟在相机和数字化仪间有变化，通常会在方形像素中发现百分之几的失真。这可以在获取图像后，通过对图像中的像素重采样，来测量并补偿这种变形，如第 4 章所述。大多数数码静物相机以 4:3（传统视频的宽高比）或 3:2（35mm 胶片的宽高比）的宽高比和正方形像素点来获取图像。

由于像素的面积有限，场景中那些跨越边界的像素会有效地平均两个区域的亮度级，并根据像素与边界的相对位置得到一个中间的亮度值。这意味着需要有较高的横向像素分辨率和大量可以分辨的灰度级来精确地定位边界。图 1.24 显示了一幅图像的几个例子，这些图像在水平方向具有不同的像素数。图 1.25 显示了同一幅图像不同灰度级的情况。

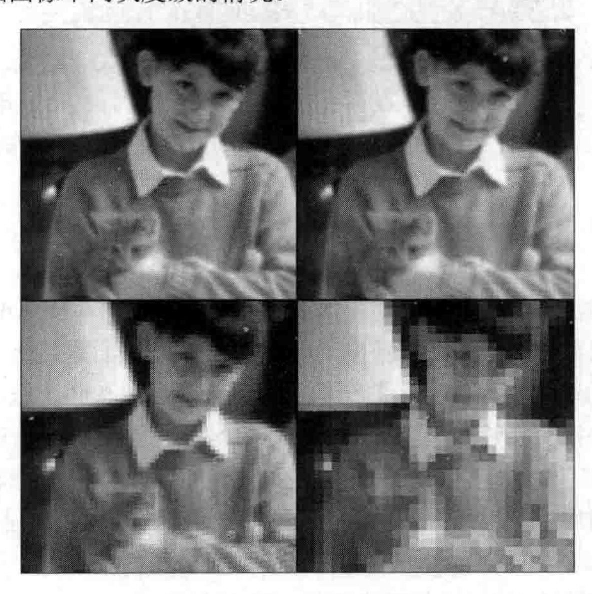


图 1.24　同一图像的 4 种表示，显示了所用像素数量的变化。从左上角起为 256×256、128×128、64×64 和 32×32。所有情形下都保留了一个完整的 256 级灰度值。平均更大像素覆盖区域的亮度，在每一步中实现了图像的粗化

对于多数常见类型的图像获取设备，如相机，像素表示跨越场景或标本的有限区域的平均信号，但也存在不属于此类的情形。例如，在低倍镜下，扫描电子显微镜束对标本中远少于图像中像素大小的体积取样。扫描探针显微镜中的探头也是如此。使用激光测距仪从 Clementine 轨道对月球的高程成像，得出了直径大约为 10cm 点的高度，但点相隔 100m 或更大。

在这些情形下，相邻像素间的关系的解释可能稍有不同。像素对离散且良好分隔的点采样，而不是跨越边界的平均。从像素到像素的中间值或逐渐变化值的情况较少，问题就变成如何定位每一边两侧的两个采样点间的边界。如果沿边界的两侧有许多点，可假定边界有一些几何形状（如局部直线），则可使用拟合方法来将它定位到像素间距的小数部分。这些方法将会在关于图像测量的第 10 章中进一步讨论。

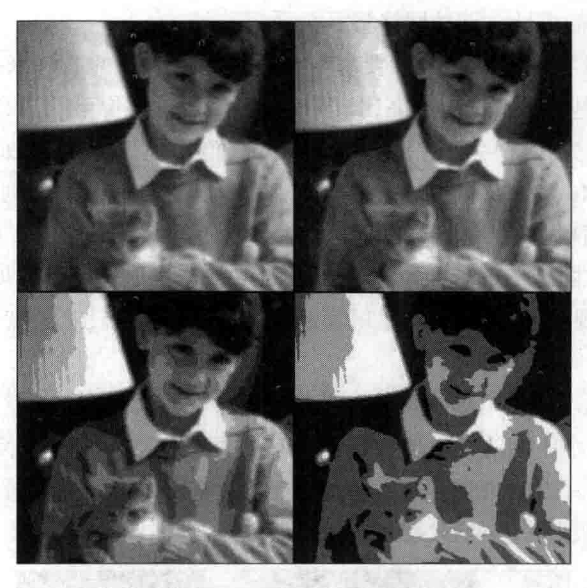


图 1.25　同一图像的 4 种表示及所用灰度级的变化。从左上角起为 32、16、8 和 4。所有情形下都保留了一个完整的 256×256 像素阵列。舍入原始像素值的亮度，在每一步中实现了图像的粗化

1.10　灰度分辨率

除了沿每个扫描行定义采样点的数量，并因此定义图像的分辨率外，ADC 的设计也控制着每次测量的精度。高速闪速模数转换器通常测量每个电压读数来产生范围在 0～255 中的一个 8 位数。这一范围可能不用于实际的图像，因为实际的图像可能不会从全黑变到全白。此外，绝大多数模拟摄像机和其他相关电子元件很少能够产生没有电子噪声的电压，从而来验证全 8 位数字化有效。具有 49dB 信噪比的典型“优良”相机意味着仅 7 位实际信息可用，第 8 位是随机噪声。但 8 位非常好地对应了计算机内存最常用的字节组织方式，因此 1 字节的存储可放入来自图像中 1 像素的亮度值。

高端数码静物相机和多数扫描仪会产生超过 256 种可分辨的亮度值，且对这些通常以 2 字节或 16 位来存储数据，从而提供了 65536:1 的可能范围，超过了任何当前成像设备（而不是可能不显示为图像的数据源，如使用扫描探针测量表面高度，详见第 15 章）的能力。对于有着 10 位或 12 位输出的相机，数值被移至最高有效位，且低位要么为零，要么为随机值。对于显示和打印目的，8 位就已足够；但对于处理与测量，更多的深度可能非常重要，后续章节将对此讨论。在许多系统中，为与更通用的 8 位范围兼容，值的直方图仍表示为 0～255，但对由浮点值组成的亮度则限制为整数。本书约定使用这种表示。

当从计算机内存中显示存储的图像时，数模转换器中使用的数字用于产生控制显示器［通常为阴极射线管（CRT）或液晶显示器（LCD）］亮度的电压。因为计算机显示技术因其他目的已发展到较高的水平，这一处理相对是无噪且高分辨率的。这些显示通常对于红、绿和蓝信号有 256 级亮度，当红、绿、蓝信号有相等的值时，可以看出结果为一种中灰值。

人眼不能完全分辨这类显示中 256 种不同的亮度级，使用喷墨或激光打印机（见第 3 章）也不能成功地记录或打印它们。CRT、LCD 或相片打印设备可分辨 20～40 个亮度级，这表明数字化仪在这方面的性能绰绰有余，或至少对于本来人眼的性能都是足够的或成像的目的就在于生成印刷品的那些应用是足够的。

得到另一种限制的稍微不同的情形出现在覆盖非常大的动态范围的图像中。真实世界场景通常包含非常明亮的区域和暗色区域。如 SEM 图像这样的科学图像中，有对应于边缘和突出部分的非常明亮的区域，以及内部暗色区域。天文图像的范围从星星非常亮的光到尘埃云或星际空间非常暗的亮度级。如果仅捉襟见肘地使用 256 个亮度级来覆盖整个范围，则对揭示明亮或暗色区域中细节的小变化不够敏感。使用高的位深如 12 位（4096 个亮度级，相当于胶片相机的性能）来获取图像时，可以记录数据，但不能成功地在显示屏或打印机上查看。方便目视解译而对这类高动态范围图像的处理方法详见第 5 章。

以非常微弱的光或某些其他成像技术如在扫描电子显微镜（SEM）中的 X 射线图所获得的图像，对图像灰度深度增加了另一种限制。当为每个图像像素收集的光子（或其他粒子）数较少时，统计波动和随机噪声会变得很重要。图 1.26 显示了高 ASA 设置（高放大增益）对图像中随机像素变化的效果。两幅图像使用相同的相机及相同的光照与光圈设置来记录；更改相机的 ASA 设置会导致不同的曝光时间。

图 1.27 显示了一幅荧光显微图像，其中的一个视频帧示出了大量的统计噪声，这一噪声妨碍了我们分辨或测量所呈现的结构。对多个帧进行平均处理会收集更多的信号，并可增强信噪比，进而提升细节的可视性。



图 1.26　用 1600（上图）和 100（下图）的 ASA 设置记录的相同图像，表明相机中较高增益产生了较大的随机噪声

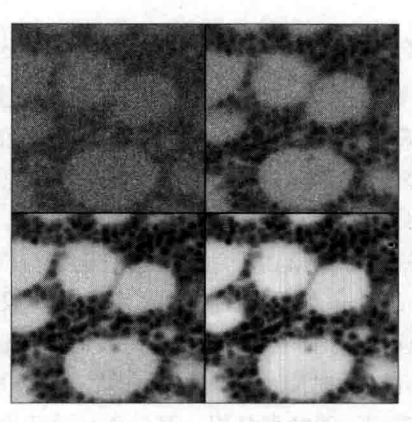


图 1.27　对带噪（低光子强度）图像（骨髓的光学显微镜图像）求平均。从左上角开始：1 帧、4 帧的平均、16 帧的平均和 256 帧的平均

1.11　噪声

在一些图像中，像素值会在原始场景理想均匀的区域内变化，其原因要么是光子或其他信号的有限的计数统计、在芯片内的电子偏移中引入的损耗；要么是放大器或电缆中出现了电子噪声。任何情形下，这种变化通常称为噪声，而由图像到噪声电平表示的场景中出现的差别所引起的对比度比值，称为信噪比。当信噪比较低时，观察者可能就看不到一些特征。图 1.28 显示了一个例子，在该例中，几个不同大小和形状的特征以不同的信噪比叠置在噪声背景上。分辨这些特征的能力通常与特征的面积成正比，而与形状无关。

在此图中，对有着最差信噪比的图像进行了平滑操作，这在一定程度上增强了特征的可视性。现有的通过图像处理来改善带噪图像的方法，将在关于空间域和频率域方法的章节中探讨。然而，如果可行，对带噪图像进行处理的最好方法是收集更多的信号并改进统计数字。

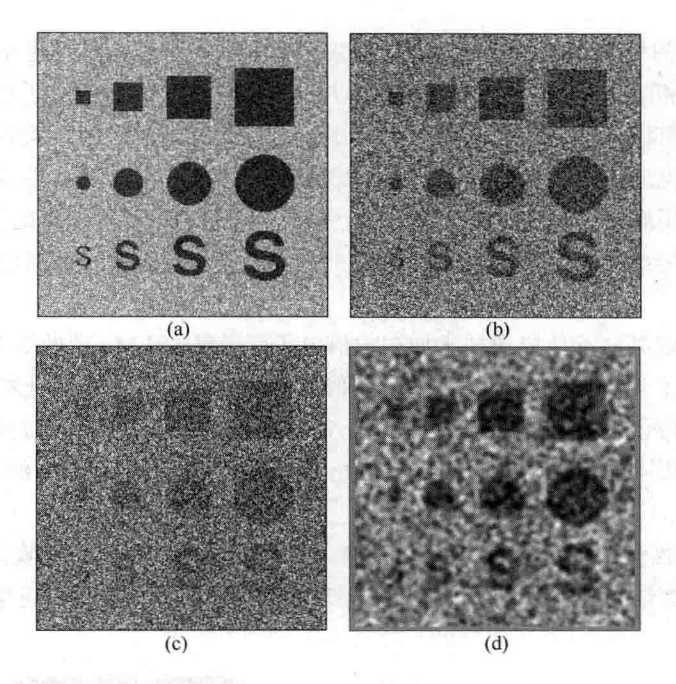


图 1.28　噪声背景上的各个特征：(a)信噪比为 1∶1；(b)信噪比为 1:3；(c)信噪比为 1:7；(d)空间平滑后的图(c)

增加曝光量（增大曝光时间、镜头光圈或光照）会因为统计效果而降低噪声，如图 1.26 和图 1.27 所示。质量的提升与所收集的光（或其他信号）的数量的平方根成正比。需要使用具有足够阱尺寸的检测器来容纳电子，并在图像中使用足够高的位深来保持对比细节。给检测器及相关的电子芯片降温可在几分钟的长获取过程中降低电子噪声。许多不带降温措施的相机芯片开始显示出令人难以接受的像素噪声，原因就在于超过几秒钟的积分时间的暗电流。

以 25～30 帧每秒的视频速率获取图像有时称为“实时”成像，当然，这一术语应正确地为足以揭示特殊应用中时间上变化的任何成像速率保留。对于某些情形，缩时摄影可能仅要求以几分钟、几小时甚至几天的时间为周期拍摄一帧。对于其他情形，则需要非常短的曝光和高速率。不使用视频帧速率（或带宽）的特殊相机，其速率可高于全帧标准视频相机的 10 倍，对于小图像尺寸，速率还可能更高。这些相机通常使用单行检测器和光学偏转（如旋转的镜子或棱镜）来覆盖图像区域。

对于许多应用，重复率并不需要这样高。频闪成像或高速快门速度可能足以定格重要的运动来提供清晰的图像。电子快门而不是机械快门可用于控制固态成像设备。可很容易达到低于 1/10000 秒的曝光时间，但这么短的曝光需要高光照亮度。

1.12　高深度图像

生成图像数据集来进行查看或测量的其他设备，会产生比相机更大范围的数据。例如，测量表面上一点的高度的探针式仪器，对于 10^5 的量程分辨率比值，其垂直精度高达几纳米，最大垂直行程几百微米。这就要求以能保留全部分辨率值的格式来存储数据，这种仪器通常使用 4 字节每像素来存储数据。从海平面到珠穆朗玛峰以 8 位（256 个值）编码的地表高程图，对应于 100 英尺/位；在这样的地形图上，佛罗里达的绝大部分地区与海平面就很难区分。使用 2 字节/像素（65536 个值）时，每位约表示 6 英寸，在地图上可以分辨沿多数街道的路肩。使用 4 字节/像素（40 亿个值），每位所对应的小于 200μm，这时甚至可记录海滩上沙子的粗糙度。

在相机有较大亮度范围的某些情形下，会以 12 位或 14 位每像素的深度存储。然而，由于深度超出了绝大多数 CRT 显示或用户查看与打印的能力，因此降低深度是合适的。如果图像的实际亮度范围未覆盖整个可能的范围，那么手工或自动缩放来选取实际使用的范围可在很大程度上降低存储需求。许多计算机软件（如 Adobe Photoshop）提供程序来从相机中读取原始线性格式的数据，以通过对光照、虚光和对比度进行调整来对其转换。

在其他情形下，尤其是在密度测量时，需要用到一个转换表。对于密度测量，期望的密度值会随亮度呈对数变化；这将在第 10 章中详细探讨。256 个亮度阶的范围并不能以有用的精度覆盖从 0 到大于 3.0 的光学密度的典型范围（即入射光照的 $1/10^3$ 被透射），因为在该范围的暗端，256 阶中的 1 个阶表示光学密度中非常大的阶。以 12 位（4096 阶中的 1 个阶）的数字化可解决这个问题，但使用对数查找表将结果值转储为一个表示光学密度的 8 位值（占据单个计算机字节）可能更为有效。

查找表（LUT）可使用硬件或软件实现。它们将原始值作为存储的索或预先计算表的索引，然后由该表提供派生值。这一处理非常快，不会影响获取过程。许多数码静物相机使用 LUT 将来自检测器和 ADC 的线性输出转换为一个值以模仿胶片的行为。这里讨论的 LUT 用于图像获取，以一个非线性表将 10 位、12 位或 14 位数字值转换为可被存储的 8 位值。LUT 也可用于显示存储的图像，特别是用颜色替代灰度值，来创建伪彩色显示，并将校正曲线应用到输出设备（显示器和打印机）来匹配色彩。这一主题将在本章的后面和第 3 章中讨论。

许多图像并未涵盖数字化仪的整个动态范围的亮度范围，其结果是图像的直方图仅涵盖所存储或显示的可用值的一部分。图 1.29 显示了这样一幅图像的直方图。图形的平坦（空白）区域表明亮端和暗端的亮度值未被图像中的任何像素用到。如图所示，通过将直方图扩展到全部可用范围来扩展亮度会增强特征的可见性，并提升局部结构的感知对比度。相同数字的亮度值从图像中丢失，如直方图中的间隙所示，但它们现在均匀分布在整个范围内。非线性拉伸直方图的其他方法将在第 5 章中介绍。

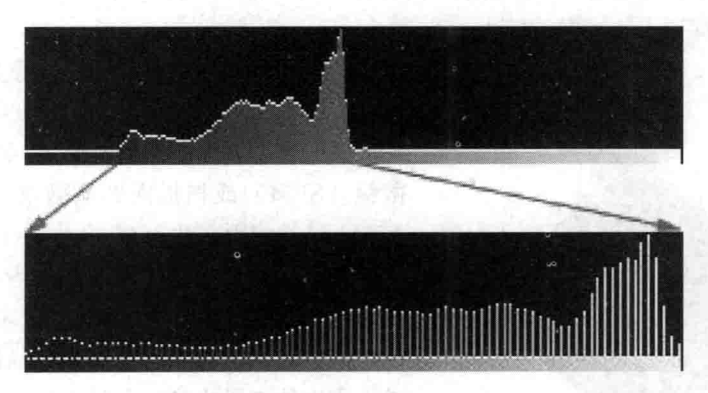


图 1.29　线性扩展直方图以覆盖存储或显示的全部范围

由于许多天文图像的对比度范围对于照相冲印而言太大，于是开发了特殊暗室技术，这些技术自推出以来一直用于其他应用与计算机软件中。“反锐化掩模”（见图 1.30）通过抑制图像的整个亮度范围，增强了显示局部对比度的能力。该抑制先稍微散焦地将一幅“掩模”图像打印到另一底片上。然后对该底片显影并放到原始图像上进行最后的打印。这种重叠会降低明亮区域的曝光度，从而显示出细节。这一方法也可用在数字图像处理中，即要么减去图像的平滑版本，要么使用拉普拉斯算子或高通滤波器（见第 5 章和第 6 章）。当存储全部深度为 12 位或 14 位的图像时，需要这一处理来在 CRT 上显示待观看的图像。

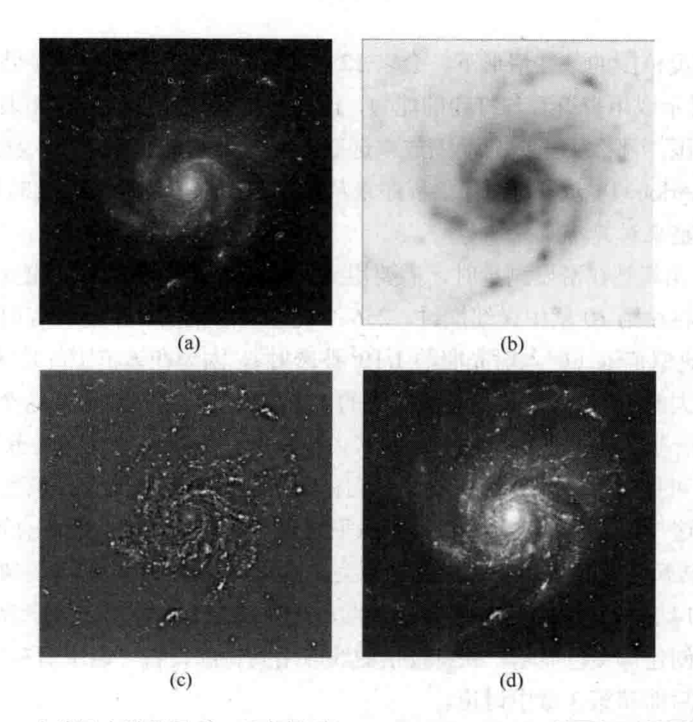


图 1.30　反锐化掩模：(a)大熊星座漩涡星系（风车星系）M101（NGC5457）的原望远镜图像；(b)“散焦”负片；(c)合并图(a)和图(b)产生了只有局部对比度的图像；(d)合并图(c)和图(a)后产生的经典反锐化掩模结果

我们需要深入理解相机的性能级别。标准视频相机有大约 30 万个传感器，且典型的数码静物相机有上千万个传感器，但人眼有大约 1.5 亿个传感细胞。此外，这些传感细胞紧密地集中于中央凹处，即注意力集中的中心区域。尽管只能分辨单个场中的数十个亮度级，但眼睛可以自动地调整到覆盖约幅度的 9 个数量级的整个亮度级，来选取最优的范围（尽管在这一范围内的最暗部分中丧失了色彩敏感性）。第 2 章将深入探讨人类视觉的能力。

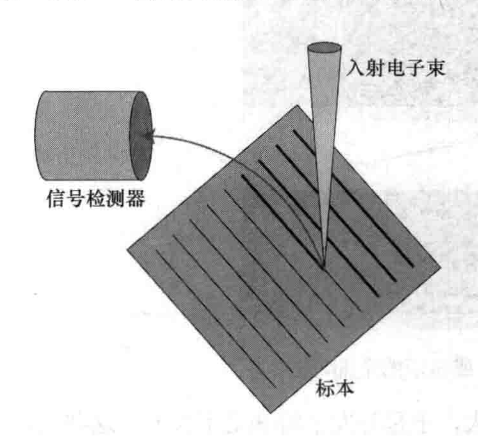


图 1.31　扫描电子显微镜（SEM）将精细的电子束聚焦在标本上，当光束以一种光栅模式扫描时，会产生可用于成像的不同信号

许多类型的科学设备不直接使用相机也会产生图像或以图像来观看的数据。比如，使用 X 射线断层术或磁共振成像仪的医学成像将在第 13 章中探讨。扫描电子显微镜（SEM）或扫描探针显微镜（如原子力显微镜或 AFM）通常使用多达 1000 个扫描行。使用 8 位或有时使用 12 位来数字化信号的设备，有着与许多相机系统类似的图像分辨率和尺寸。图 1.31 示意地表明了 SEM 的功能。聚焦的电子束以光栅模式扫描样本的表面，同时通过检测出的电子来产生各种信号。这包括次级电子和背向散射电子、X 射线、可见光和样本中的电子效应。

其他点扫描显微镜，如 AFM、共焦扫描光学显微镜（CSLM）甚至接触式表面光度仪等，会产生非常不同的信号与信息。通过获取扫描速度和参数，它们都提供与样本上空间位置相关的时变信号，进而将数据存储为图像。许多这样的设备在平行与垂直扫描方向上，有着不同的噪声与分辨率特性。

由平板扫描仪可生成较大的像素阵列。这些设备使用一个线性固态检测器阵列，通常至少可扫描 8 英寸×10 英寸的面积，有时甚至可达这一面积的几倍。尽管主要用于桌面出版市场，但它们也可用

于扫描平的或近似平的许多类型的标本，以及用于蛋白质分离的电泳凝胶，或照相冲印或底片。高质量底片可记录几千个可辨识的亮度级以及每英寸几千个点（这两个值都远大于打印件）。扫描仪也用于对诸如医学 X 射线和 35mm 幻灯片这样的底片进行数字化。商用扫描仪可用于出版业，将胶片转换为数字值以便于存储（例如以柯达的 Photo-CD 格式）。在消费级别，2400 像素/英寸的扫描仪通常适用于大幅面反射原图，而高达 4000 像素/英寸的扫描仪则适用于 35mm 幻灯胶片。绝大多数这样的扫描仪可数字化全彩 RGB 图像。

扫描仪并不昂贵，但速度非常慢，需要花数十秒钟的时间才能对扫描区域数字化。它们所表现出的另一典型问题是传感器中的模式噪声：如果一行中的所有检测器的性能不一致，则在传感器扫描图片时，会在图像中出现“条带”效应。如果扫描运动不完全平滑，则会在其他方向产生条带。到目前为止，这种扫描仪最大的问题是在光照方面。在靠近扫描区域边缘的位置，亮度会下降，原因在于光源的长度较短。更麻烦的是，光源需要预热一段时间才能变得稳定（与荧光照明相比，LED 的问题较少）。

1.13 彩色成像

绝大多数现实世界的图像是彩色而非单色的。光学显微镜产生彩色图像，许多生物标本制备技术会使用色彩以新的方式来辨识结构或定位化学活动。甚至对于无机材料，使用偏振光或表面氧化会产生彩色图像来描述其结构。SEM 通常是一种严格的单色成像工具，但 X 射线能量或反射电子能量可能会引入颜色。图 1.32 显示了来自用 SEM 成像的一个矿物样品的 X 射线信号的单个灰度图像。X 射线图像显示了 9 种元素（Al/铝，Ca/钙，Fe/铁，K/钾，Mg/镁，Na/钠，O/氧，Si/硅，Ti/钛）中每种样本

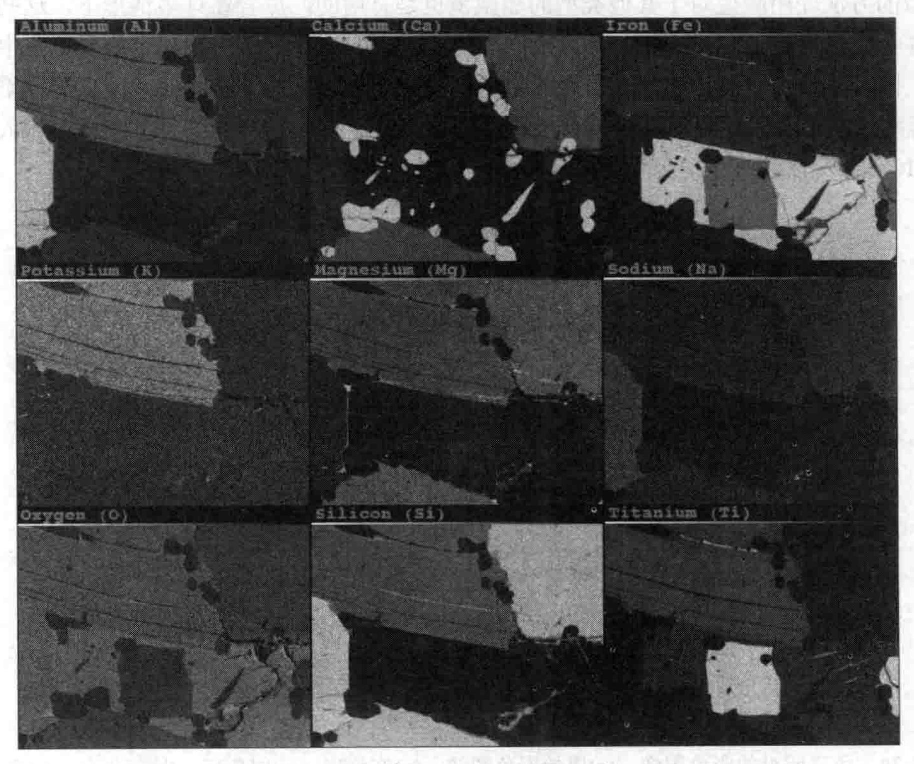


图 1.32　抛光云母样本的扫描电子显微镜图像。各幅 X 射线图像或“点图”显示了样本中各种矿物质内对应元素的位置

的成分变化。通常会有更多的图像，而不只是红色、绿色和蓝色显示通道，并且没有明显“正确”的选择将颜色分配到元素。图 1.33 给出了一些可能性，但要记住的是，单幅彩色图像不可能一次显示所有的元素。任意指定几种（如到显示的红色、绿色和蓝色平面），可帮助用户判断包含两种或多种元素的区域或面积的对齐。第 5 章介绍将处理多通道图像的方法，以找到主要成分。现在彩色 X 射线图在 SEM 很常见，这些图形类似于使用离子微探针测量强度得到的浓度图，但其他颜色的使用，如透射电子显微镜（TEM）中能量的损耗，仍在实验中。

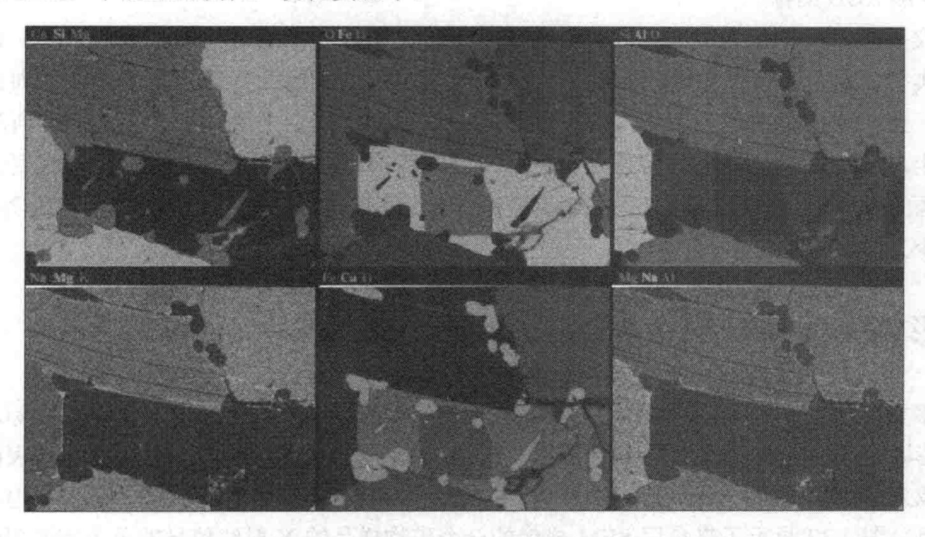


图 1.33　将图 1.32 中不同图像指定为红色、绿色或蓝色通道而形成的混合彩色图像，它有助于勾画出各种矿物质

用类似的方法使用颜色对于其他类型的显微镜也是有用的，尽管在许多情形下，这些可能性还未在商用设备中找到。对于宏观成像工具，也同样如此。一个简单的例子是使用色彩在空中交通控制中表示高度。使用色彩增加了传递多维信息的带宽，但有效使用这些方法需要对用户进行培训并可能得益于一些标准化。当然，还要求用户具有全彩色视觉能力。图 1.34 显示了不同程度色盲人士（男性中约有 5%～10%的人会是色盲，女性则相对较少）对色彩的观感。

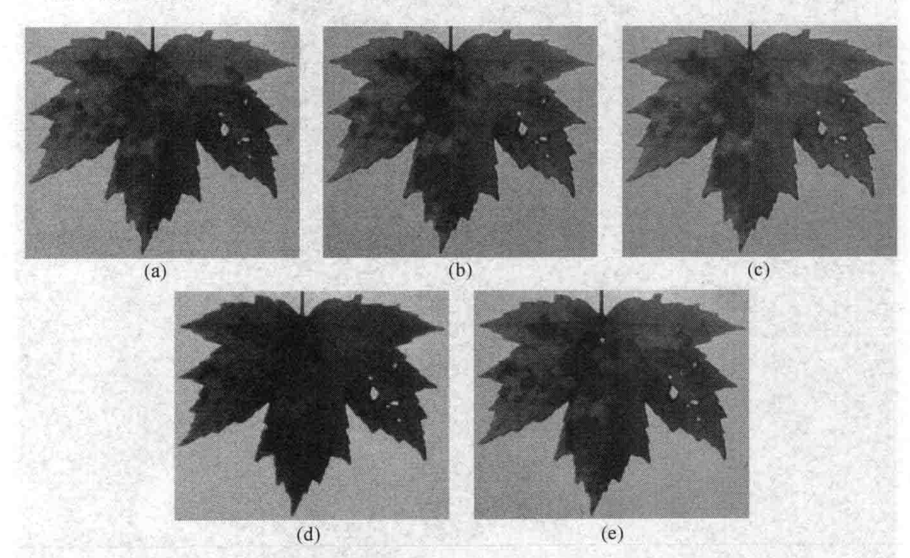


图 1.34　色盲所见的彩色图像：(a)原始彩色图像；(b)红色盲；(c)绿色盲；(d)蓝色盲；(e)全色盲

使用色彩来对多维信息进行丰富的编码，必须与单色图像中常用的假彩色或伪彩色来替代颜色区

分开。使用伪彩色的原因在于视觉能力分辨亮度细微差别能力的限制。尽管仅可分辨单色图像中 20～40 个灰度阶，但却可察觉出几百种不同的彩色。此外，彩色有助于人们将感兴趣的特征描述为“暗红－橙色的”而不是“中度灰色的”。

使用彩色温标替代亮度值，可让我们显示并看到很小的局部变化，并全局地识别图像中的相同亮度值。由于这是下面将要探讨的成像的目的，因此会非常有用。伪彩色一直特别用于空间探测器返回的许多图像中。知道有多少人认为土星环是色彩鲜艳的或哈雷彗星事实上被七彩光晕环绕是很有趣的。使用伪彩色的危险在于，它会掩盖图像的真实内容。色彩会强迫我们将注意力集中于图像的细节而丧失完全形态或整体性。本书中的图像处理例子选择性地使用伪彩色来示例一些处理效果及所产生的像素值的变化，但通常伪彩色会使得人眼的注意力不集中于所增强的图像内容。

对每个离散存储的亮度值，此处使用伪彩色显示以替代来自存储的颜色或预先计算表格中的颜色。如图 1.35 所示，这应与用彩色显示来识别结构或指出特征属性的其他用途区分开。这些也依赖于使用彩色来向观察者表达丰富的信息，但要求在这种信息可用之前，对图像进行相当多的处理与测量。

色彩通常用于对表面高程着色（见第 15 章）。在科学可视化中，色彩用于速度、密度、温度、成分和许多其他并不明显的属性。这些属性通常与图像的特性无关，只是简单地利用人的能力来区分比灰度值更多的色彩。

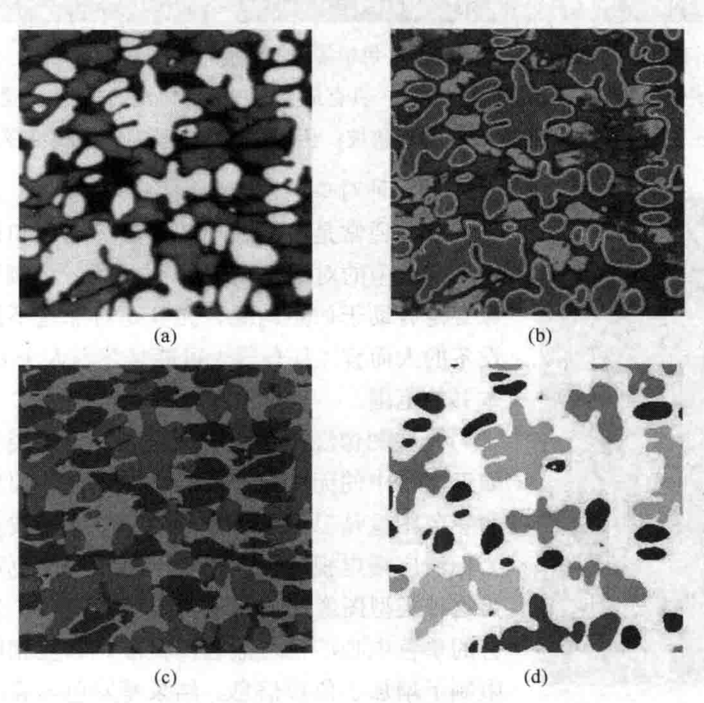


图 1.35　伪彩色显示的不同用途：(a)抛光金相样本显微灰度图像的一部分，它具有不同平均亮度的三个相；(b)使用彩色替代灰度值的伪彩色调色板或 LUT 图像（注意沿亮相与暗相之间边界出现的误导颜色）；(c)把色彩分配到各个相的图像。此时需要对图像进行阈值处理或基于灰度亮度和相邻像素分类，将每个像素指定给一个相来分割图像；(d)原始图像中最亮的特征已根据特征大小分配颜色。此时需要执行创建图(c)的步骤，然后将所有相接像素组合到特征中，进而测量特征

多数基于计算机的成像系统很容易就可将所存储图像中的亮度值替换为色彩的各种查找表（LUT）。它和前面所述的输入查找表的工作方式一样。存储的灰度值用于选择 LUT 中的一组红色、绿色和蓝色亮度，LUT 控制发送给显示管的电压。许多系统还提供实用工具来创建这些彩色表，但很少

提供指南来帮助构建有用的表。因此必须要小心，且通常需要包含一个渐变来显示已用于特殊图像的彩色温标。一种方法是沿一条穿过色彩空间的路径系统且逐渐地改变颜色。图 1.36 中显示了几种不同颜色表的例子。渐变可帮助我们组织场景的不同部分。快速转变颜色增强了梯度并易于查看局部变化，但可能会完全掩盖图像中的一些完整内容，如图中的最后一个示例所示。

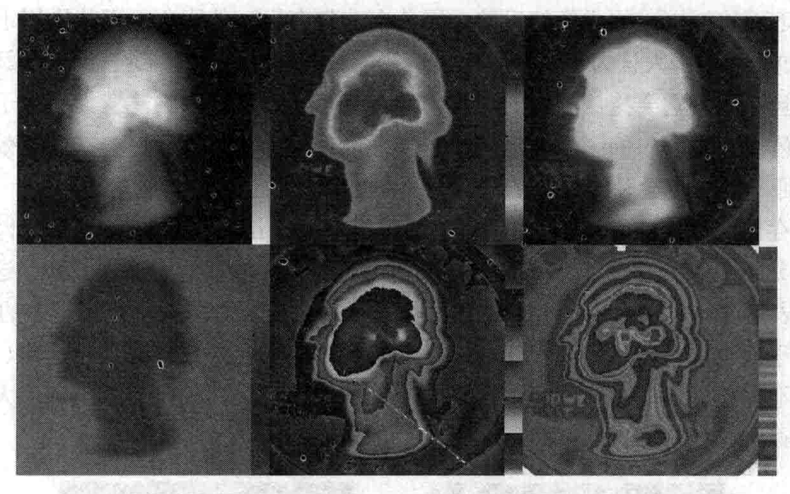


图 1.36 显示查找表（LUT）的几个例子，每幅图像都带有色阶。从左上角起：原始单色或灰度图像；谱或彩虹（色调的变化，具有最大饱和度和恒定的亮度）；暖色阶；三原色的三色混合；六种饱和色的亮度斜坡；色调随饱和度和亮度的线性变化做正弦变化

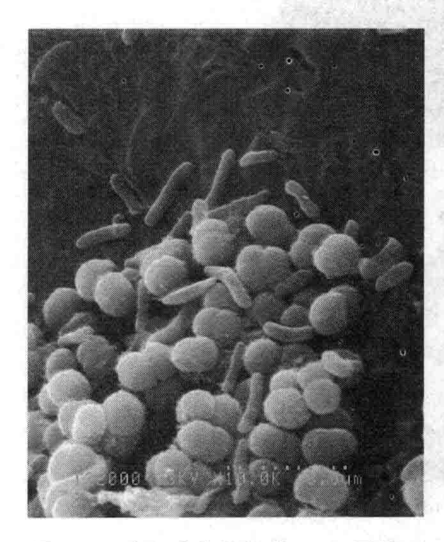


图 1.37 人工着色的细菌 SEM 图像，其中已应用色彩浓淡来从背景中区分杆状细菌和球形球菌

另一种对单色图像使用色彩的应用，是特征的手工上色或着色。这通常是根据人类观察者的理解和识别来完成的，以区分不同类型的对象。图 1.37 显示了一个例子。当然，这样的图像必定有助于信息沟通，尤其是对那些不熟悉物体或操作成像设备的人而言（且有些人可能是艺术人士），但此类过程超出了本书的范围。

一些图像源使用色彩来对各种不同类型的信息进行编码，如天文学中的无线电波的强度和极化。但到目前为止，最常见的彩色图像是以记录可见光的三个不同波长的亮度产生的。因为标清广播电视使用在视觉上可接受的色彩，因此视频足以作为这种类型图像的介质。美国使用 NTSC 彩色编码方案作为已有的单色电视广播可兼容的扩展。它在相同的已经很窄的带宽限制下增加了色彩信息。结果是彩色与亮度信息相比有着较小的横向分辨率。

使用比光强（亮度）数据低的分辨率来表示彩色数据的彩色子采样，可降低图像中的数据量。本章后面介绍的 YUV 色彩空间（Y 表示光强或亮度，U 和 V 表示色度或色彩）中就是这样做的。未压缩的 YUV 色彩表示为 4:4:4。常见的子采样选项是 4:2:2、4:2:0、4:1:1 和 YUV-9。

4:2:2。对 Y 全带宽采样并对 U 及 V 的 2:1 水平采样。这是专业和广播视频及磁带格式（如 D-1 和数字 Betacam）中最常使用的采样方案。它看上去不错，但数据压缩率仅为 33%。

4:2:0。对 Y 全带宽采样并对 U 及 V 在水平与垂直两个方向上以 2:1 采样。也就是说，对于每 4

个光强样本，每隔一行有两个色度样本。这会使数据量缩减 50%。4:2:0 是大多数 DVD 采用的 MPEG 压缩所使用的色彩空间。

4:1:1。对 Y 全带宽采样并对 U 及 V 以 4:1 水平采样。这是许多数字摄像机中使用的数字视频(DV)格式的色彩空间。它以 4 个像素宽和 1 个像素高应用 U 和 V 样本，所以在水平方向的渗色要比垂直方向的更差。

YUV-9。这是互联网上绝大多数视频压缩使用的色彩格式。对于一个 4×4 的像素块中每 16 个光强 Y 样本，仅有一个 U 和一个 V 样本，产生具有更低色彩保真度的更小文件。YUV-9 子采样通常会在明亮色彩物体的边缘处产生可察觉的彩色伪影，尤其是红色。

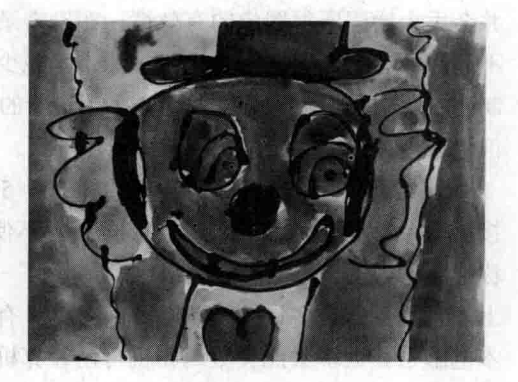
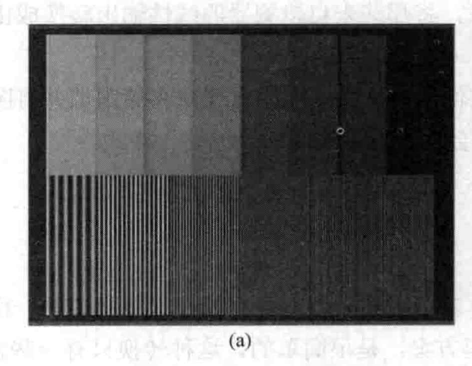


图 1.38　儿童小丑画。所绘区域并不恰好以线为边界。暗线表明了特征的尺寸和形状，观看者不会被超出区域的颜色混淆

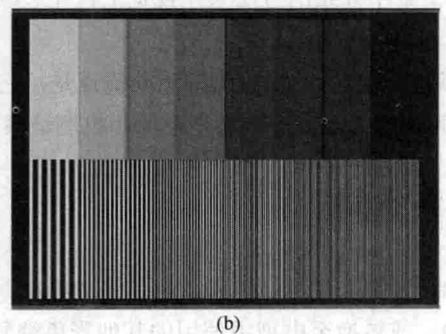
这些限制对于电视图片来说是可接受的，因为观看者能容忍并不是那么陡然划界的色彩，并使用由图像中并不一定完全对应的亮度分量定义的特征的边缘。同样的容差也被画师有效利用，对其孩子还未学会在线条范围内涂色的父母而言也是熟悉的。图 1.38 给出了一个例子，其中跨界的渗色或区域内的变化并不会使眼睛感到混淆。

NTSC 彩色的较差空间清晰度与其表示实际彩色值时的较差一致性是匹配的［一个常见的笑话是使用 NTSC 表示“Never The Same Color”（颜色从来不对）而非“National Television Systems Committee”（美国国家电视系统委员会）］。与单色图像的录制相比，彩色图像的磁带录制对于分析更没价值。但广播信道强加的限制并不一定意味着相机和其他分量会没用。Super-VHS 或 S-video 录制设备（也称为分量视频）可增强这些图像中彩色信息的清晰度，在这些设备中，亮度或光强及色彩或色度信息是单独传输与录制的，不对数据进行子采样。

用于技术目的的一些彩色相机会独自产生红色、绿色和蓝色信号，从而单独进行数字化。在计算机内存中记录图像时，要将每个信号视为一个单色信号进行处理，将它转换为一组数字，并存储在内存中。如果信号开始已组合在一起，则使用的编码方案可能是 YIQ 或 YUV（在后面定义），它与 NTSC 广播方案紧密相关。不将色彩和亮度信息混在一起可更好地保留图像的保真度。替代在一行上传输复合信号，一些相机和录像机会将色度（色彩）和光强（亮度）信号分开到不同的行上。这种所谓的 Y-C 或 S-Video 格式用于某些摄录机中（Hi-8 和 S-VHS 格式）。许多计算机接口接受该格式，它明显改善了数字图像的质量，如图 1.39 所示。



(a)



(b)

图 1.39　复合视频信号和分量视频信号的比较：(a)复合视频信号；(b)信号生成器产生的分量视频信号，它是使用相同的接口板生成的。注意黑色条纹与白色条纹分辨率的不同及不同颜色之间边界的不同

视频领域的另一项重要进展是数字视频（DV）录制。与模拟卡带式录像机一样，数字视频使用一个活动的磁头，当磁带移过磁头时，通过旋转磁头的方式将每个扫描行以某个角度写到磁带上。信号被编码为一组数字比特，因此具有许多优点。就像 CD 技术取代模拟音频磁带一样，数字视频信号并不会在图像的传输或副本制作过程中受到保真度损失的影响。更为重要的是，由于传统广播电视强加的 4MHz 上边频限制，在模拟录制中被抛弃的高频信息在 DV 中得到保留。数字视频录制高达 13.5MHz 的照度（亮度）信息及其 1/4 的色度（彩色）信息。这会产生更清晰的边缘，尤其是图像中的亮度数据，并会大大增强所得图像的有效性。在磁带录制时，数字视频信号得到完全保留，这和传统模拟录像机不同，后者会对数据强加进一步的带宽减少，且很容易受到由于磁头变脏、磁带磨损或从母盘不断复制副本所引起的质量下降。绝大多数更新的 DV 摄像机使用（固态）闪存取代盒式磁带来存储数据，但仍采用相同的数据格式。

因此 DV 图像的像素分辨率约为 500×500（但色彩信息更少，如上面提到的 4:1:1 格式），对比度接近 8 位，因为它们已经是数字格式，不使用单独的数字化仪就可读入计算机中。用于数字视频（也称为固件）的 IEEE 1394 标准协议建立了一个标准的串行接口标准，存在于消费级价位的相机和卡座上，且受到了许多计算机制造商的支持。有一些廉价的接口板，它们内置在一些个人计算机甚至笔记本电脑中。大多数此类数码相机可由计算机控制来选择将要传输的单个帧。它们也可用于记录标注有日期和时间的帧，将数字视频盒式磁带或闪存芯片转换成很小但保真度非常高的存储格式来存储几百幅单独的图像。

1.14 数码相机的局限性

几乎所有传统的胶片相机制造和摄影用品供应商，都参与了消费市场数码静止帧相机的当前开发。因为光学（有着几何失真的定焦镜头）的局限性和有限的分辨率（甚至移动电话中的低端相机芯片的分辨率也要比传统的模拟视频录制的分辨率高），其中的许多并不适合技术应用。有些相机会在芯片中对像素进行内插，进而创建空放大的图像。在多数情况下，这会在图像中产生严重影响定量测量的缺陷。

但对图像进行压缩会产生更严重的问题。要使图像所占的内存最少，可使用 JPEG 和其他形式的图像压缩。第 3 章中解释了这些技术是有损技术，它们会丢弃图像中的信息。数据丢弃时是有选择的，以便对人眼解译与识别熟悉图像的影响最小，但对图像量化分析的影响可能很大。边缘被断开并移位，色彩和密度值被更改，细节可能会被消除或移动。许多消费级相机也在相机内部对图像进行处理，例如，使用第 5 章中介绍的方法来在视觉上锐化边缘，将那些来自检测器的线性输出转换成由胶片提供的对数响应。

更高端消费级和专业级相机的重要优点是，它们能以无损压缩的方式来存储图像并将图像传送给计算机，因此非常受人们的欢迎。不同相机制造商会使用许多不同的“原始”格式。

1.15 色彩空间

从 RGB（各个红色、绿色和蓝色信号的亮度，就像由相机获取并存储在计算机中的一样）转换至 YUV 或 YIQ，或转换至电视中使用的其他彩色编码方案，是很简单的，这种转换只有一些舍入误差，不会丢失信息（但在彩色信号使用有限的带宽广播时，会出现损耗）。“光强”信号 Y 只是全色的亮度，如黑白电视接收机所显示的单色图像。它按人眼对其敏感的比例对红色、绿色和蓝色信号进行组合。

彩色信号的 I 和 Q 或 U 和 V 分量被选择用来与广播中使用的硬件兼容。如图 1.40 所示，V 信号本质上是红减青，而 U 是品红减绿。YUV 或 YIQ 和 RGB 的关系见式（1.1）和式（1.2）。将编码的 YUV 或 YIQ 信号转换为 RGB，是通过矩阵转置来实现的，如下所示。

$$
\begin{aligned}
Y &= 0.299R + 0.587G + 0.114B & R &= 1.000Y + 0.956I + 0.621Q \\
I &= 0.596R - 0.274G - 0.322B & G &= 1.000Y - 0.272I - 0.647Q \\
Q &= 0.211R - 0.523G + 0.312B & B &= 1.000Y - 1.106I + 1.703Q
\end{aligned} \tag{1.1}
$$

$$
\begin{aligned}
Y &= 0.299R + 0.587G + 0.114B & R &= 1.000Y + 0.000U + 1.140V \\
U &= -0.147R - 0.289G + 0.436B & G &= 1.000Y - 0.395U - 0.581V \\
V &= 0.615R - 0.515G - 0.100B & B &= 1.000Y + 2.032U + 0.000V
\end{aligned} \tag{1.2}
$$

RGB（和打印所用的互补 CMY 减三原色）和 YIQ 及 YUV 都是针对硬件的方案。RGB 源自相机传感器和显示荧光粉的工作方式，而 YIQ 和 YUV 则源自广播考虑。图 1.41 显示了由 RGB 信号定义的"空间"：它是一个笛卡儿立方体空间，因为红、绿和蓝信号彼此无关，且可以相加来在立方体内生成任何颜色。存在对图像处理更为有用且与人类感知更接近的其他编码方案。

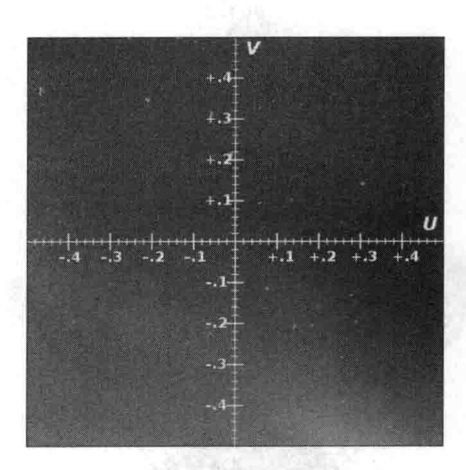


图 1.40　UV 色彩空间。V 轴由红色变化至青色，U 轴从品红色变化至绿色

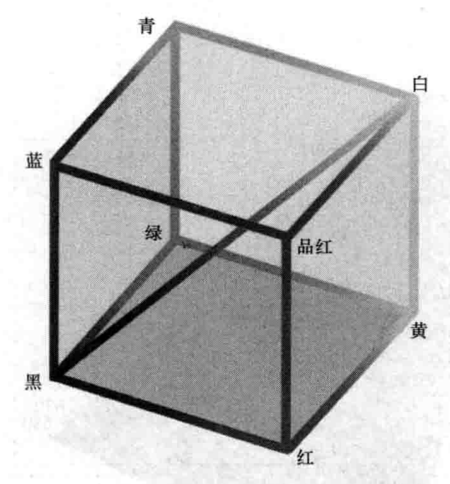


图 1.41　RGB 色彩空间。红色和绿色结合产生黄色，绿色加蓝色产生青色，蓝色加红色产生品红色。等比例的红色、绿色和蓝色沿立方体对角线相结合，产生从黑色变化至白色的灰度。青色、黄色和品红色是在印刷中使用的减法三原色，它们减去白色（纸张色）后分别成为红色、蓝色和绿色

CIE（Commission Internationale de L'Eclairage，国际照明委员会）色度图是一幅定义色彩的二维曲线图，如图 1.42 所示。第三个（垂直）轴是光强，它对应于全色的亮度，像 YUV 中的 Y 值一样，生成一幅单色（灰度）图像。所显示的两个主要彩色轴称为 x 轴和 y 轴，它们总是正的，组合在一起可定义可见的任何颜色。

用于彩色度量的仪器使用 CIE 三原色来定义任何颜色的主要波长和纯度。混合任何两种颜色，对应于沿两种原始颜色间的一条直线选取图形中的一个新点。这意味着 CIE 图上一个三角形的三个角在用于 CRT 中发射荧光粉或 LCD 面板中使用的滤波器的红、绿和蓝位置，这个三角形定义了显示器可显示的所有颜色。有些颜色并不能通过混合这三个荧光色来创建，原因在于它们位于三角形之外。用于显示器或任何其他输出设备的颜色范围称为色域；印刷设备的色域通常比显示器的小，详见第 3 章的讨论。图中有界区域的边缘对应于纯饱和色，并用单位为纳米的波长来标示。

过中心点（对应于白光）画一条直线，可以在 CIE 图中显示补色。因此得到一条从绿色开始、穿过白色再到品红色的直线。CIE 图的缺点之一是，它不能指出人眼能察觉的颜色的变化。有时，可通过在图中画一系列椭圆来证明。在小变化很难察觉的绿色区域，这些远比其他地方大。饱和度的变化（从中心到边缘的距离）通常比色调（围绕图的位置）更容易察觉。

CIE 图提供了一种定义色彩的方式，但它既不对应于硬件的操作，也不直接对应于人类视觉。一种有效的方法体现在 HSV（色调、饱和度和纯度）、HSI（色调、饱和度和亮度）及 HLS（色调、明度和饱和度）坐标系中。它们彼此紧密相关，且与艺术家对色泽、明暗和色调的感知一致。在这一系统中，色调就是多数人认为的颜色，例如红和黄之间的区别。饱和度是所呈现的颜色数量，例如红和品红间的区别。第三个轴（称为明度、亮度或纯度）是光量，即暗红和淡红之间或深灰和浅灰间的不同。

可以看出，画出这三个值的空间是一个圆或一个六边形锥体或双锥体，有时可显示为一个球体或柱体。将该空间画成一个双锥体最为有用，在双锥体中，锥体的轴是从黑到白的灰度变化，距中心轴的距离是饱和度，方向是色调，三原色 RGB 颜色相隔 120°。图 1.43 示意说明了这一概念。

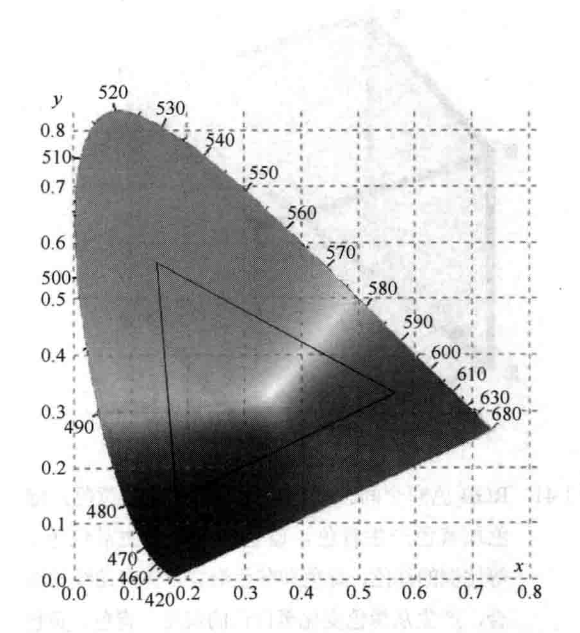


图 1.42　CIE 色度图。沿边缘出现的色彩是完全饱和的。数字给出了以纳米为单位的光的波长。内接三角形表明典型彩色 CRT 显示器的色彩可通过混合荧光粉中的红光、绿光和蓝光来产生

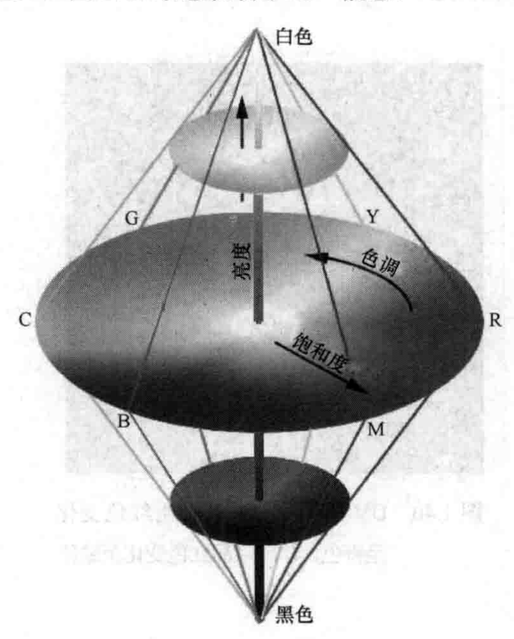


图 1.43　色调－饱和度－强度空间的双锥表示。灰色沿中心轴。到轴的距离给出了饱和度，而方向（角度）表明色调

对于图像处理与理解色彩来说，该空间有许多优点。例如，如果第 4 章中给出的算法（如空间平滑或中值滤波）被用于降低一幅图像中的噪声，那么将它们分别应用到 RGB 信号会导致结果中的颜色偏移，但将它们应用到 HSI 分量中则不会产生这一结果。此外，在分割（见第 7 章）过程中使用色调来区分特征通常对应于人类感知并忽略了阴影效应。另一方面，由于 HSI 分量并不对应于多数硬件的工作方式（既不对应于获取也不对应于显示），因此要求进行计算来将使用 RGB 编码的图像转换为 HSI 或从 HSI 转换回 RGB。

RGB 空间和色调－饱和度－亮度坐标间的转换有多种方式，具体取决于所用 HSI 空间的形状。最常用的选择是球体、圆柱体或双锥体。在所有情形下，强度轴均对应于 RGB 立方体的主对角线

（图 1.41 中的灰度线），但所有 HSI 空间几何都不完全符合该立方体的形状。这意味着要在两个空间中表示颜色，在转换过程中饱和度值必须有一定程度的失真。在后面的图 1.61 所示例子中可看出这一点，其中一幅高分辨率全色卫星图像用于替代由三个低分辨率 R、G、B 图像组合构成的图像中的强度通道。

HSI 空间对图像处理是很有用的，因为它们会以对应于人类视觉系统响应的方式来分开彩色信息，同时也因为它们的轴对应于标本的许多物理特性。这样的一个例子是生物组织的染色。对于有用的近似而言，色调表示所染的颜色，饱和度表示所染的数量，亮度表示标本的密度。类似地，在遥感图像中，色调通常用于识别土地使用情况（农作物、水体和城市），而亮度会随当地太阳照射角度而变化。但这些空间在数学上是不方便的：不仅因为色调值循环通过从 0° 到 360° 的角度并折回，而且圆锥形空间意味着增大亮度或照度会改变饱和度。

对大多数应用而言，与 HSI 方法足够接近且在数学上更易于处理的、几何上更简单的空间是球状 Lab 模型。L 是灰度轴或光强，而 a 和 b 是组合在一起来定义颜色和饱和度的两个正交轴（见图 1.44）。a 轴从红（$+a$）至绿（$-a$），b 轴从黄（$+b$）至蓝（$-b$）。在该空间中，色调并没有像在 HSI 空间中所示色轮那样的相同角度分布。相反，它们基于如下思想：没有像红－绿或蓝－黄这样的颜色。20 世纪 60 年代中期的发现表明，相对而言，彩色一定程度上与人类视觉系统相关，视网膜颜色刺激被转化为亮与暗、红与绿及蓝与黄之间的差别［尽管 Crane and Piantanida（1983）已证明在有些情形下，可以感觉到红－绿和黄－蓝的混合色］。L、a 和 b 轴在对应于硬件的 RGB 空间的正交简单性与在许多彩色管理系统、分光光度计和色度计中使用的更加基于生理的空间（如 HSI）之间，提供了一个实用的折中。L、a 和 b 及 R、G 和 B 间的数学关系如式（1.3）所示。

$$\begin{bmatrix} L \\ a \\ b \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{-\sqrt{2}}{6} & \frac{-\sqrt{2}}{6} & \frac{-\sqrt{2}}{6} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \quad \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ 1 & \frac{-1}{2} & \frac{1}{3} \\ 1 & 0 & \frac{-2}{3} \end{bmatrix} \begin{bmatrix} L \\ a \\ b \end{bmatrix} \tag{1.3}$$

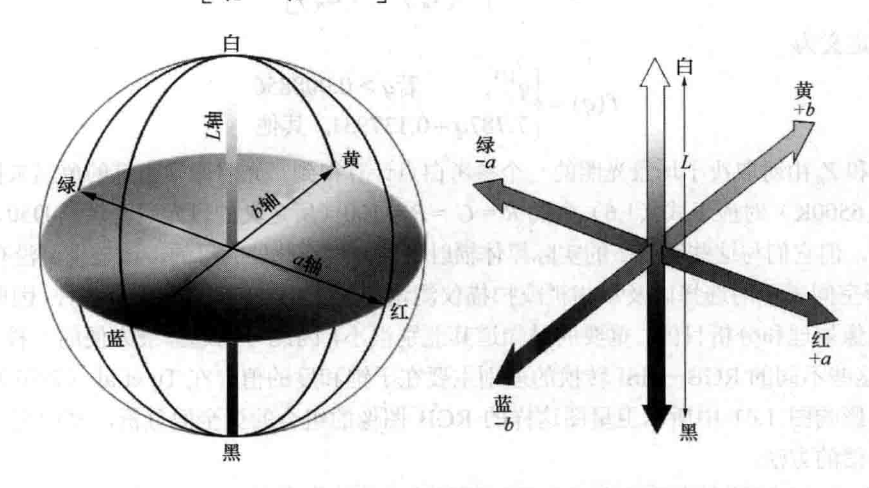


图 1.44 Lab 色彩空间。灰度值沿南北向垂直通过球体的线分布。相同的 a 轴和 b 轴也用在更精确的 CIELab 空间中

由这些值可以看出，在需要处理或分析时，色相和饱和度可被计算为

$$\begin{aligned} H &= \arctan(b/a) \\ S &= \sqrt{a^2 + b^2} \end{aligned} \tag{1.4}$$

另外，可以直接进行从 RGB 至圆锥 HSI 空间的转换：

$$
I=\frac{R+G+B}{3}
$$

$$
S=1-\frac{3\cdot\min(R,G,B)}{(R+G+B)}
$$

$$
H=\begin{cases}\arccos(z) & 若G\geqslant R\\ 2\pi-\arccos(z) & 若G\leqslant R\end{cases} \tag{1.5}
$$

$$
z=\frac{(2B-G-R)}{2\sqrt{(B-G)^2+\dfrac{(B-R)}{(G-R)}}}
$$

CIELab 色彩空间是另一种失真的几何，它被认为是“感觉上均匀的”，这意味着一种刚好能检测到的视觉差别由空间中任何位置或方向的一个常数距离表示。尽管以下公式中有着明显的 6 位数值精度，因为它基于有限数量的人类测试者，因此在一定程度上有点过于简单，且从技术上讲，这一空间应用于特定光照条件下的印刷物检查，而不能像计算机显示器那样地发射显示。但 CIELab 还是用做比较色彩的标准空间。它有几个缺点，最严重的缺点是，简单的径向直线并不能保持常数色调，且需要定义一个白点。从 RGB 到 CIELab 的变换需要有一个中间步骤，称为 *XYZ*:

$$
\begin{aligned}
X&=0.412453R+0.357580G+0.180423B\\
Y&=0.212671R+0.715160G+0.072169B\\
Z&=0.019334R+0.119193G+0.950227B
\end{aligned} \tag{1.6}
$$

基于这些 *XYZ* 值，L^*、a^*和 b^*分量为

$$
\begin{aligned}
L^*&=116\cdot f\left(\frac{Y}{Y_n}\right)-16\\
a^*&=500\cdot\left[f\left(\frac{X}{X_n}\right)-f\left(\frac{Y}{Y_n}\right)\right]\\
b^*&=200\cdot\left[f\left(\frac{Y}{Y_n}\right)-f\left(\frac{Z}{Z_n}\right)\right]
\end{aligned} \tag{1.7}
$$

式中，函数 f 定义为

$$
f(q)=\begin{cases}q^{1/3}, & 若q>0.008856\\ 7.787q+0.137931, & 其他\end{cases} \tag{1.8}
$$

式中，X_n、Y_n 和 Z_n 由对取决于场景光照的一个参考白点计算得到。光照通常由灯的色温来描述，如标准的 D65 亮度（6500K）对应于式（1.6）中的 $R=G=B=100$（所定义的日光照明颜色 D50、D65 等通常称为黑体温度，但它们与这些温度下的实际黑体辐射不同，对于此处的目的，该差别倒没有什么关系）。

由于色彩空间表示的选择以及从相机或扫描仪测量的 RGB 值转换的方法有多种，因此对于后续章节中描述的图像处理和分析目的，重要的是知道其差别很小，因此可以选择最方便的一种。Ledley et al.（1990）证明这些不同的 RGB－HSI 转换的差别主要在于饱和度的值。在 Tu et al.（2001）中可以找到这些差别如何影响图 1.61 中所示卫星图这样的 RGB 图像的组合的更全面分析，文中提出了对饱和度的变化进行补偿的方法。

将 Lab 或 HSI 空间的亮度修改为对应不同颜色的感知响应也是可能的。用于此目的的典型模型如式（1.9）所示。所示的其他权重因子有着很高的精度，如（$0.212671R+0.715160G+0.072169B$），它最初是从电视显像管中使用的特定荧光粉推导出来的，用于将彩色视频信号转换成单色，甚至已正确地应用到了线性（非伽马校正）强度，但在许多情形下用于表示不同彩色通道的近似作用。在真实情形下，实际权重取决于特殊相机中所用的彩色滤镜，它也会因相机的不同而不同，同时也取决于光照条件。

$$
I（或L）=0.25R+0.65G+0.10B \tag{1.9}
$$

注意，红、绿和蓝的加权组合并不会产生视觉上中性的灰色。其目的是，对各种颜色进行加权，

以便所组合强度的均匀步长（比率）产生可察觉亮点增量的均匀集合。相等的红色、绿色和蓝色亮度会在计算机屏幕上产生视觉上无色彩的灰色。

用于数字化彩色图像的硬件要么接收直接来自相机的 RGB 信号或 Y-C 分量信号，要么接收一个复合信号（如 NTSC）并使用电子滤波器将单独的分量分开，进而提取红、绿和蓝信号。对于上面讨论过的单色情形，这些信号接下来被数字化以产生数值，对于 *R*、*G* 和 *B*，这通常是 0～255 范围内的 8 位值。它占用 3 字节来存储每个像素，所以 640×480 像素的图像要占用计算机中近 1MB 的存储空间。对于大多数摄像机和电子设备，信号中并不包含这一丰富的信息，且最低的 2（或更多）位是噪声。因此，一些系统仅为 *R*、*G* 和 *B* 保留 5 位，它可放到两个字节中。这一降低对于互联网图形和桌面出版应用而言，通常是足够的，但按这种方式处理时，就难以提取颜色信息来进行处理或分析操作。

彩色图像通常数字化为 24 位 RGB，这意味着对红、绿和蓝存储 8 位或 256 个（线性）亮度级。这足以允许在荧光屏或计算机屏幕上显示，或进行打印，但取决于数据的动态范围，这可能不足以测量图像内的较小变化。由于摄影胶片可获取较大的动态范围，用于透明度或胶片扫描的一些扫描仪提供更大的动态范围，对于 *R*、*G* 和 *B*，它通常是 12 位（4096 个线性级别）。专业质量的数码静物相机在其“原始”图像中也能获取这一数量的数据。当图像存储到计算机中时，这些 36 位图像可被约简为一个“最优”的每通道 8 位。转换色彩或灰度存储的原始动态范围的问题之一是，由固态检测器测量的亮度值是线性的，而胶片和人类视觉近似是对数的，所以在图像的暗区，可被存储的最小亮度阶相对较大，且可能导致视觉伪影或密度值的较差测量精度。

使用查找表指定最好的 256 种颜色来匹配图像的内容，可将彩色图像进一步降低到 256 色，它可用于计算机图像中来降低文件大小。得到的图像通常称为“索引色”图像。查找表本身仅需要 3×256 字节来为 256 种颜色指定 *R*、*G*、*B* 值，它们被写到了显示器硬件中。仅需要从该调色板中选取每像素单个字节。选取最优调色板的最通用技术是中位切分算法（Heckbert，1982），它基于原始图像中像素的实际 RGB 值来细分色彩空间。在一些应用中，默认的颜色查找表是预先定义的（且对于互联网浏览器而言，仅包含 216 个标准色）。对于视觉目的，这样的降低通常会提供令人满意的显示图像；但对于图像分析目的，则应该避免这种降低。

许多情形下，尽管获取了彩色图像，但“绝对的”色彩信息对图像分析没有用。相反，它是从一个区域到另一个区域的相对颜色差别，可用于区分所呈现的结构和物体。此时，可能不需要使用全彩色图像。存在不同类型的色彩分离方法。分离 RGB 分量（或互补 CMY 值）通常用于桌面出版中，但不常用于图像分析。也可以将图像分离成色调、饱和度和强度分量（通常称为通道）。图 1.45 给出了一个例子。某些结构在一个通道中可能比在另一个通道中更明显（例如，白百合花瓣上的粉红色斑点在红色通道中不可见，但在绿色或饱和度通道中却非常明显）。选择一种合适的色彩空间可简化某些应用；例如，如前所述，在光学显微镜下观察的染色后的生物组织中，色调通道显示了染色的位置，饱和度通道显示了染色的数量，而强度通道则记录了标本的密度。

隔开图像中任何数量的特殊“颜色”（色调）是可能的。这一计算（技术上是感兴趣彩色向量与图像中每个像素的点积）等效于在单色相机前面物理地插入一个透射滤光片。该滤光片可选择性地吸收补色（例如，蓝色滤镜通过吸收黄光来使黄色区域变暗），并透射同一蓝色的光，因此所得图像的对比度可基于图像中的彩色分布。例如，摄影师一直使用黄色滤镜来使蓝天变暗，并为云彩产生大的对比度的单色图像。图 1.46 是一个例子，其中计算机用来应用滤镜。选取一种颜色来最大化所呈现结构之间的对比度的颜色，将产生一幅其亮度适合进行阈值处理的图像（详见第 7 章）。因为可指定任何想要的色调，并且可用一幅存储的图像进行操作，应用计算机的彩色滤镜与应用物理滤镜相比，灵活性通常更大。

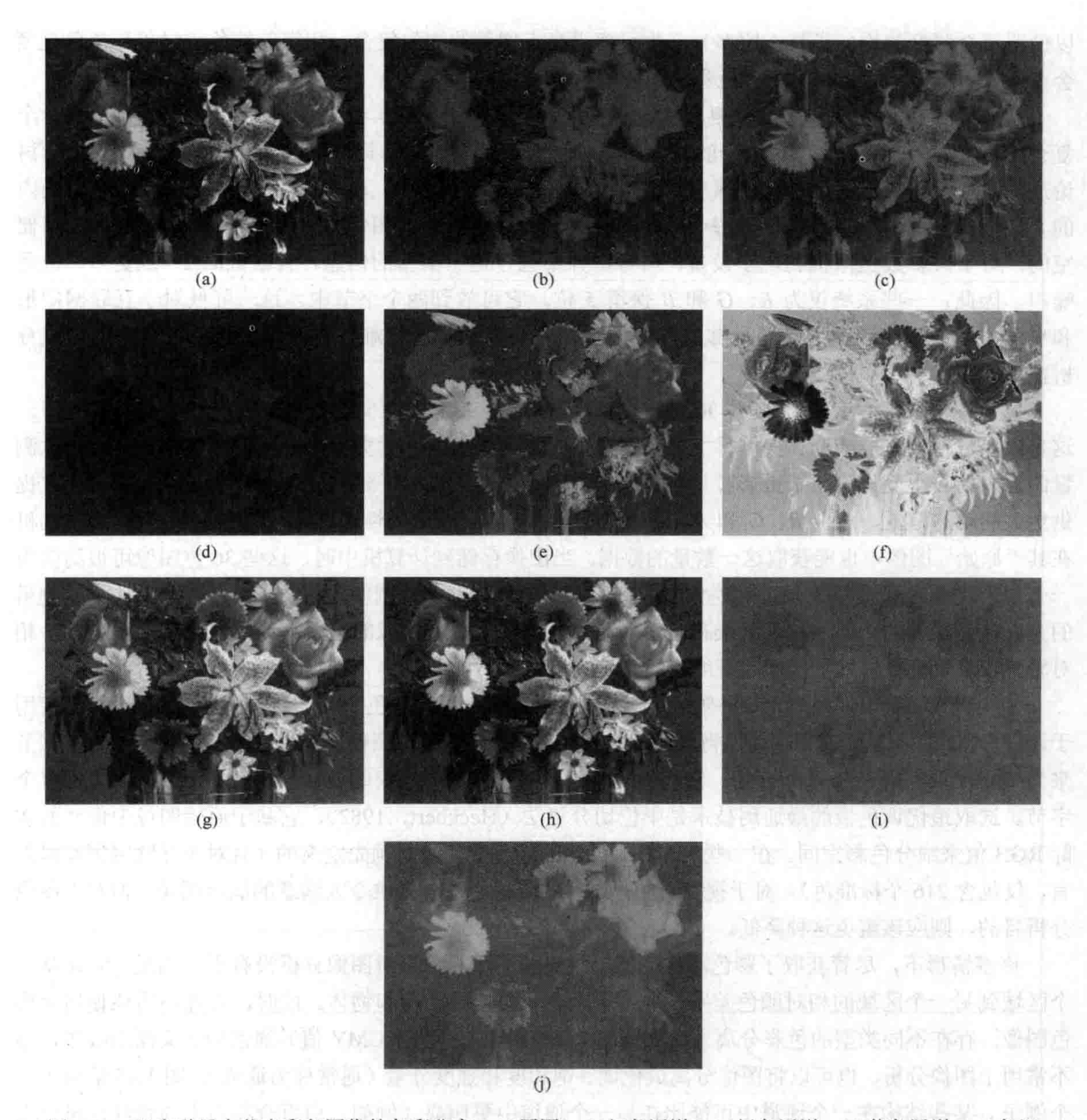


图 1.45　现实世界中花卉彩色图像的色彩分离：(a)原图；(b)红色通道；(c)绿色通道；(d)蓝色通道；(e)色调通道；(f)饱和度通道；(g)强度通道；(h)L 通道；(i)a 通道；(j)b 通道。注意，由于权重的不同而得到不同的波长，所以在 HSI 表示法的图(g)中的强度与在 Lab 表示法的图(h)中的 L 分量是不一样的

在许多情形下，将一幅彩色图像化简为单色灰度图像是很有用的，除了选择一个 R、G、B、H、S、I、L、a 或 b 通道或应用一个彩色滤镜的明显方法外，还可使用许多其他方法来实现。如果使用所期望得到图像中呈现的所有结构之间的最大灰度对比度来简化灰度图像的阈值处理和测量，则可为每幅图像计算一个唯一的函数，该函数将拟合一条线，这条线通过那些表示色彩空间中所有像素的彩色坐标点。这条最小二乘拟合线给出了各个像素彩色值的最大分离度，且投影到该直线上的每个像素的坐标位置可用做一个灰度值来提供最优的对比度（Russ，1995e）。图 1.46(d)给出了一个例子，注意图 1.46(c)中单个彩色滤镜结果为所选的结构提供了更好的对比度，而回归法最大化了图像的整体对比度。

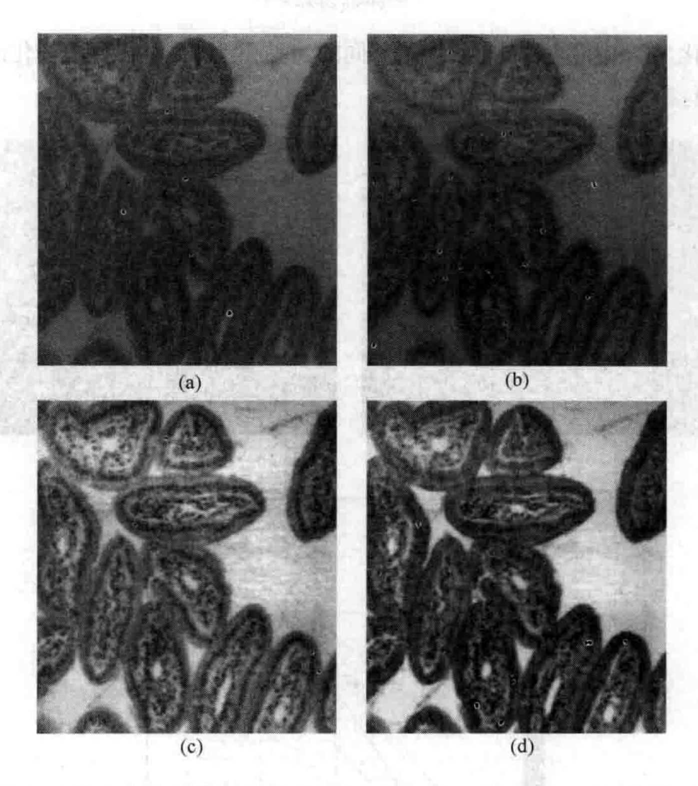


图 1.46 嵌入老鼠小肠的石蜡的显微薄片机组织切片，用 Feulgen 染色和固绿色 Bouin 反向染色：(a)原始彩色图像；(b)对原始彩色图像应用 480nm 蓝色滤镜后的结果；(c)图(b)中的单色强度；(d)回归计算的最优灰度

1.16 色彩校正

在不同的光照条件下获取图像后，记录的彩色值会受到影响。人类视觉能容忍光照相当大的变化，明显的方法是使用视场的外围来归一化色彩解译，或假设某些已识别出的对象具有已知颜色。如第 2 章中指出的那样，人类视觉系统的特性也会产生一些明显的彩色错觉。

一些数字视频和静物相机都有一个自动的白点校正，这一校正允许记录一幅无色彩的灰色卡片图像，并使用它来调整后续的颜色。图像获取后，可在软件中进行类似的校正，方法是调整红、绿和蓝的相对数量，进而设置一个已知的没有色彩而只有纯灰色的区域。如图 1.47 所示，这一方法很简单，但仅适用于有着对应于白色、黑色和一种或多种中性灰色区域的图像。

扩展这一方法可生成存储好的校准曲线，从而匹配几种设备的颜色。使用包含许多已知颜色的色卡来进行调整，可在相机或扫描仪、各种 CRT 和 LCD 显示器、各种类打印机甚至投影仪之间产生彩色匹配。为给获取设备构建校准曲线（当然，包括所用照度的特性），典型的过程是对一个已知的彩色阵列进行摄影或扫描。由一个存储文件中打印一个类似阵列后，使用分光光度计测量每种颜色，可为打印机构建类似的曲线（如第 3 章中所述，这对于所使用的油墨及纸张是特定的）。在屏幕上显示颜色并使用分光光度计测量它们会产生另一条校准曲线，它用于调整发送到显示器的值。

其结果是视觉上颜色到处匹配的图像。市场上已有一些这样的产品，其中之一（GretagMacbeth Eye-1）的组成如下：可用于发射显示器（CRT、LCD）和反光（印好的）硬拷贝的一个分光光度计、一个色卡（见图 1.48），以及尽量不引人注意地执行必要操作和计算的适当软件。由于 CRT 中的荧光及 LCD 显示

器中使用的背光的老化，打印机在不同时候使用不同的油墨和纸张，且光照的变化改变相机的色彩响应，因此需要经常重复进行校正。

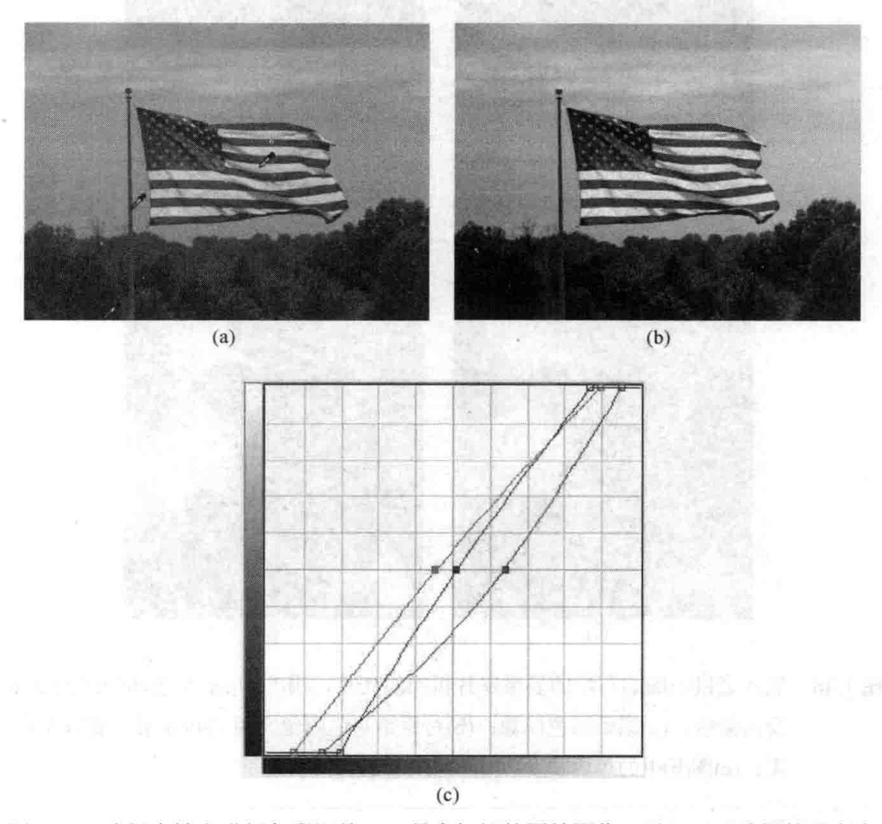


图 1.47 选择中性点进行色彩调整：(a)具有标记的原始图像，显示了已选择的具有中性暗值、中灰色度值和亮值的三个点；(b)调整后的结果；(c)红色、绿色和蓝色色彩调整显示了三个选定点的原始值（横轴）和产生新（纵轴）的曲线

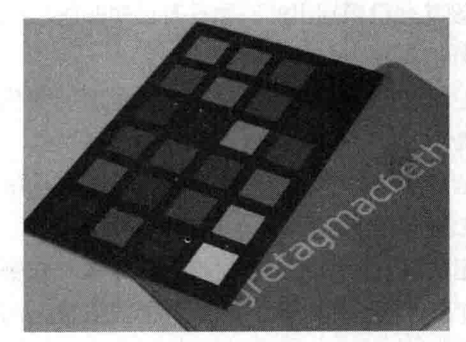


图 1.48 The GretagMacbeth ColorChecker 色卡。书中未充分地再现该色卡的颜色，本图只起说明性作用。颜色从左上角起分别为：暗肤色，亮肤色，蓝天，树叶，蓝花，蓝绿色，橙色，紫蓝色，中等红，紫色，黄绿色，橘黄色，蓝色，绿色，红色，黄色，品红，青色，白色，中性 8，中性 6.5，中性 5，中性 3.5 和黑色

对图像获取进行色彩调整的一种类似但更简单的方法是三色刺激校正。这要求对具有纯红色、纯绿色和纯蓝色区域的测试图像进行测量，如图 1.49 中显示的校准卡。实际情形下，这样的一张卡可置于被记录的场景中。因为传感器上方的彩色滤镜会覆盖波长中较宽且重叠的范围，相机（或扫描仪）记录每个区域中的红色、绿色和蓝色强度。这会产生一个值的矩阵，如表 1.1 所示，它可被求逆来产生三色刺激校正矩阵。对于图像中的每个像素，通过以测量到的 R、G、B 值作为次数乘以该矩阵来计算新的红色、绿色和蓝色值，可产生一幅其校准区域已调整到纯红色、纯绿色和纯蓝色的图像，图像中的其他颜色也被校正，如图所示。

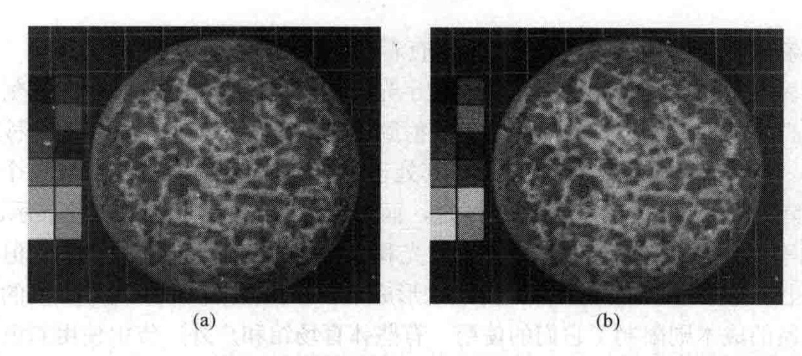


图 1.49　使用 RGB 值进行三色刺激校正：(a)原始图像；(b)采用表 1.1 计算的三色刺激矩阵校正后的图像

表 1.1　使用三色刺激值的彩色校正

从将被校正为纯红色、纯绿色和纯蓝色的三个目标区域测量 RGB 强度。在矩阵中输入强度：				
测量的强度		红	绿	蓝
区域	红	131.91	43.08	69.16
	绿	72.91	152.66	53.74
	蓝	56.34	63.74	125.98
中间的归一化强度矩阵（上面的值除以 255）				
		0.5173	0.1689	0.2712
		0.2859	0.5987	0.2107
		0.2209	0.2500	0.4940
Excel（MINVERSE）函数计算的逆矩阵				
通道		红	绿	蓝
结果	红	2.6066	−0.1680	−1.3593
	绿	−1.0154	2.0978	−0.3375
	蓝	−0.6520	−0.9863	2.8027

对于显微图像，通过在幻灯片上沉淀已知的颜料来构建一组合适的色靶（Hoffmann et al.，2005）。不一定需要纯红色、纯绿色或纯蓝色。只要三种颜色已知并在色彩空间中分得较开，就可计算校正矩阵。

1.17　彩色显示器

当然，作为将信息传递到计算机用户的载体，色彩也是很重要的。许多计算机使用彩色阴极射线管（CRT）显示器，与电视机相比，它有更高的分辨率，但基本操作原理相同。更小的荧光点、更高的扫描频率和更高的带宽放大器及单个逐行扫描（而不是隔行扫描），会产生更高的清晰度和色彩纯度。

笔记本电脑和台式计算机的平板显示器使用液晶显示器（LCD）。有源 LCD 显示器使用单独的晶体管来控制每个 RGB 单元，它们一起组成一个像素。由于更低的亮度和更窄的视场角，有些彩色 LCD 比 CRT 差。对于每种颜色，它们也有着较差的亮度和对比度范围，进而减少了可被显示的不同颜色数目。色彩及白色的强度取决于背光的颜色，而背光可由荧光管或发光二极管产生。

LCD 也用于某些投影仪中来在屏幕上显示图像。一种常见的设计使用非常亮（且色彩已校正）的光，以及较小的 LCD 面板和合适的光学器件。高亮度投影仪通常使用带有红、绿和蓝滤镜的三个单独的单色 CRT。对于这些投影仪，分辨率通常较高，因为每个 CRT 都有连续的荧光粉涂层。但是，需要仔细对齐光学元件，以保证三幅图像配准；每次设备移动或甚至加热时，都需要重新调整。在室内光

照控制不理想的条件下，如何得到足够的亮度来查看大图像也是一大挑战。

另一类显示器使用由德州仪器公司开发的数字光调制原理。使用硅晶圆上的光刻产生的小镜子阵列，用于将来自光源的光通过适当的光学器件反射到观察屏幕上。镜子可以在几纳秒之内在“开”和“关”位置翻转。来回快速移动每个镜子可控制其处于“开”位置的时间，并控制每个像素的亮度。旋转的滤镜轴允许阵列顺序地投影红、绿和蓝通道，而眼睛则将其感知为一种彩色显示。

其他类型的平板计算机显示器，包括电致发光和等离子，基本上都是单色的，但可通过滤波产生彩色图像。原理上可排列红、绿和蓝 LED 阵列来形成平板显示器，但产生足够蓝光的困难性和到目前为止这种器件高昂的成本则限制了它们的使用。有些体育场馆和户外广告中使用彩色光阵列来显示图像。纽约时报广场和东京银座就是典型的例子，但全彩色有源显示器已非常常见了。

用三基色荧光粉（RGB）的组合可产生在 CRT 上显示的彩色范围。每种荧光粉的亮度是通过调制 CRT 中撞击每种荧光粉的电子束的强度来控制的。对每种颜色使用单独的电子枪并将彩色点排列为三元组是实现这一控制的最常用方法。为阻止杂散电子撞击相邻的荧光粉，可为每个点的三元组使用带有小孔的金属荫罩。三个电子束中的每一个通过荫罩上的同一小孔并撞击相应的点。荫罩会增大图像的清晰度和对比度，但降低了 CRT 产生的光的总强度。

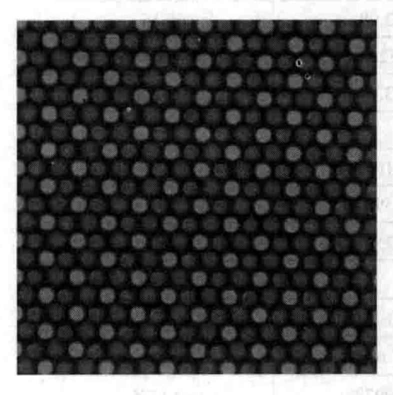


图 1.50　CRT 显示器中红色、绿色和蓝色荧光粉光点的三元组

增加亮度的一种更简单的设计可在 CRT 的垂直条纹上应用荧光粉色彩。它要么使用荫罩上的槽形图案，要么根本不使用模板。索尼特丽珑（Trinitron）设计的简单性使得显像管的成本更低，在垂直方向上镜片没有曲率，但具有高的显示亮度，以及较少的对齐问题。但是，荧光粉及电子束的垂直扩张有模糊屏幕垂直方向边缘的趋势。尽管这种设计在家用电视中相当常见，但绝大多数高性能的计算机 CRT 则使用荧光粉的三元组（见图 1.50），因为它会使图像的清晰度更高，特别是对于线与边缘。

在 LCD 显示器中，红、绿和蓝滤镜的三元组通常也排列为平行的垂直条纹［见第 3 章中的图 3.1(b)］。高分辨率计算机显示器，不管是 CRT 还是 LCD，都有一个小于 200μm 的间距（从一个三元组到下一个三元组的距离）。在正常的观看距离处，单个彩色点或条纹在视觉上并不能分辨，从而产生连续图像。

1.18　图像类型

在人眼视觉上所感受到的日常图像中，一个点上的亮度是亮度和位置及所观察表面的方向和性质的函数。由于它们的三维特性、光照与各种表面的交互以及一些表面遮掩其他表面的事实，“表面”和“真实世界”图像通常难以使用计算机算法来解译。即使是对于那些其中的优先性不是问题的相对平的场景且光源得到良好控制或至少已知的情况（如卫星成像），其表面方向、颜色、材质和其他变量的组合，会使得独立定量地解译这些参数变得困难。对于一个仔细准备好的平坦表面（如在典型的入射光显微镜中），将颜色和对比度解译为描绘粒子、纹理、内含物或其他结构则更为成功。对于图像形成及光照和光学元件的作用的其他背景知识，详见 Jahne（1997）。

第二类显微图像显示了穿过样本的光（或其他辐射）的强度［关于光学显微镜及其成像用途，详见 Smith（1990）、Bracegirdle & Bradbury（1995）、Bradbury & Bracegirdle（1998）和 Heath（2005）等］。透射图像从一个通常已知亮度和颜色的均匀光源开始。在每个点处光的吸收是样本沿该路径的密度的测度。对于某些类型的透射图像，如使用电子和 X 射线形成的图像，也可能会出现由样本中原子或分

子结构形成的辐射的相干散射所导致的衍射效应。这些会使得分析更加复杂，因为衍射与晶格或其他周期结构严密相关。

为说明简单密度之外的因素引起的复杂化，图 1.51 显示了薄钴箔的透射电子显微（TEM）图像。明显的结构是在该铁磁材料中的磁畴。在每个条带形的磁畴中，围绕原子旋转的电子已自发地对齐了。原子结构、样本密度或厚度没有变化，但图像当然会使得观察者认为出现了这样的变化。

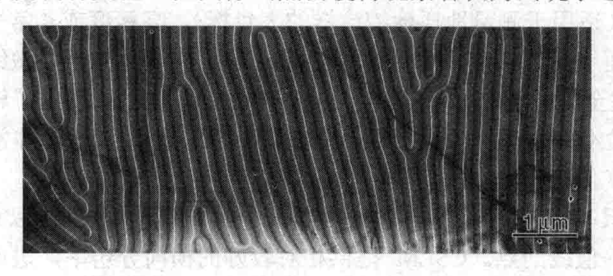


图 1.51　薄钴箔的 TEM 图像。条纹图案揭示了磁畴，其中原子的电子自旋在两个可能的方向之一上对齐

同样，图 1.52 显示了以高分辨率 TEM 查看无定形碳胶片上胶体金粒子的图像。所谓的原子分辨率表明衬底上点的图案多少有点随机，但在金粒子内部，它们是规则排列的。这种排列是粒子的晶状结构的结果，且这些点与原子位置相关。但是，这些点不是简单的原子；结构和图像间的关系是复杂的，且基本上取决于显微镜的参数以及镜头的散焦量。从预测的结构来计算期望的图像对比度是可能的，对于简单的结构可以用常规方式实现。逆运算（由图像计算结构）更为有趣，但只能通过迭代法完成。

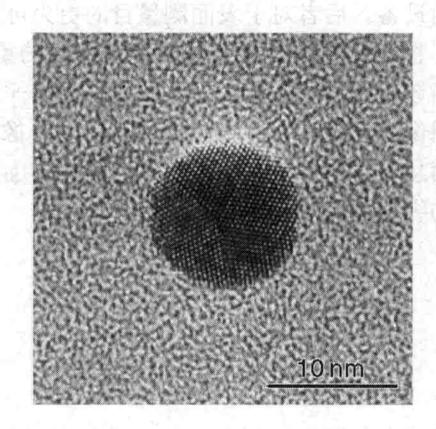


图 1.52　无定形碳基底上的一个胶体金粒子的 TEM 图像，用于显示很高的显微镜分辨率

在其他类型的透射图像中，辐射的某些颜色（或能量）可能被样本按照其化学成分或存在的选择性着色或染色。有时，这些颜料自身也会发出不同颜色的光，这些光可用成像来对局部的特殊结构定位。原理上，这与扫描电子显微（SEM）中产生的 X 射线图（见图 1.32）非常相似，在 SEM 中，电子会激励样本的原子，进而发射它们的特征 X 射线。在许多这种类型的发射图像中，标本密度及厚度的变化，或其他元素或成分的出现，至少会导致在根据所选目标元素或成分的浓度或数量来解译像素亮度时出现一些困难。

航拍和卫星图像会显示相对平坦的表面（至少是场景与相机的距离相比），因而使测量简化。尽管不是直接控制的，但光照应用到整个场景（对云彩、坡度的局部变化等有一些限制）。以各种波长波段（包括红外线）获取图像，可帮助选择与识别地表上的结构、农作物和土壤。

许多类型的工业质量控制图像也会用于测量与分析。照相机和光源的受控放置，简化了图像的解译。各种成分的大小与位置的测量可能不完全是理想二维的，但至少总是相同的。

1.19　距离成像

另一类重要的图像使用像素亮度来记录距离。例如，表面的原子力显微图像显示了表示为灰度（或伪彩色）值的表面上每个点的高程。距离图像由光栅扫描显微镜生成，如扫描隧道（STM）显微镜和

原子力（AFM）显微镜，或由使用表面光度仪触针的物理扫描生成。它们也可通过干涉光学显微镜生成，或用激光测距、合成孔径雷达、侧扫声呐和其他技术在更大的尺度产生。以像素亮度值表示距离信息的图像也可通过其他方式间接得到，如体视法、从明暗恢复形状及运动模糊。

可用于非常平坦表面测量的大多数工具会在 x、y 光栅或网格中的每个点处提供单值读数。第 15 章详细介绍了表面距离图像的处理与测量。这种数据集无法分辨可能出现的底切。就像雷达和声呐的波长范围在厘米至米那样（适用于测量地质地貌这样的大目标），高精度或非常平坦的表面，需要更小的测量尺度。对原始的表面或穿过它们的垂直截面使用 SEM、传统的光学显微镜或共焦扫描光学显微镜（CSLM），并不能令人完全满意。SEM 的横向分辨率非常好，但由于使用了较窄的孔径角，因此深度分辨率较低。难以执行立体对的测量，且计算表面的高程图或距离图像是非常耗时的，所得的深度分辨率与横向分辨率相比，仍相当差。

传统的光学显微镜具有比 1μm 更好的横向分辨率，但景深既未大到足以查看整个粗糙的表面，且没有足够浅以孤立沿着等值线的点。CSLM 提供相对较好的横向分辨率，进而大大降低了景深，同时拒绝了失焦位置的散射光。其结果是通过垂向移动样本并在每个位置保持最亮的光值来对整个粗糙表面成像，或通过在对图像中的每个点获得最亮的反射光值时，跟踪样本的垂向运动来产生一幅距离图像的设备。后者对于表面测量目的更为可取。在所有方向上，其分辨率等是 1μm 甚至更好。

图 1.53 显示了脆性陶瓷断裂表面的重构视图。它是由 26 个在 z 方向间距为 1μm 的平面形成的，仅当该位置亮于任何其他平面时，每个平面记录一个像素。透视图可旋转，以呈现出表面的更为理想的图像。但是，以距离图像（其中每个像素的亮度对应于被记录最亮反射的平面）的形式来画出相同的信息对于测量而言更为有用。图 1.54 显示了相同表面的这种表示，以及沿着穿过该表面的一条任意线的纵剖面。

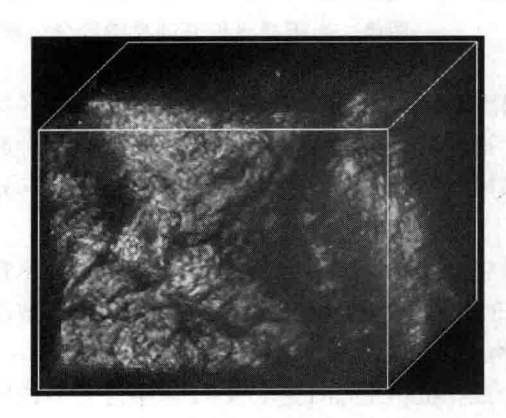


图 1.53 陶瓷的脆性断裂表面的重建 3D 图像，用共焦扫描光学显微镜成像

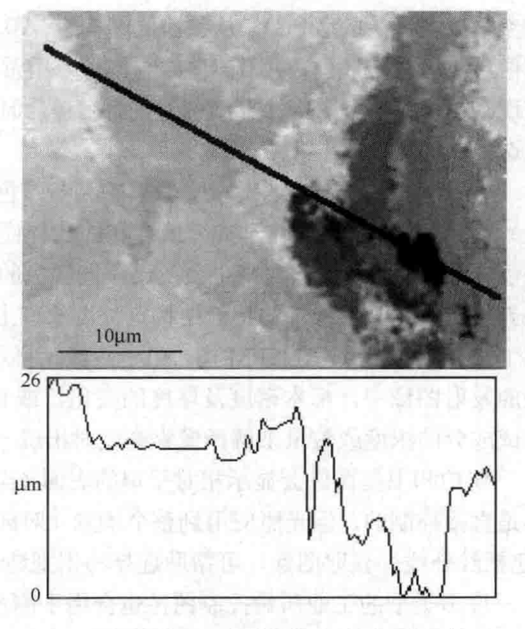


图 1.54 从图 1.53 中的数据产生的距离图像，以及沿一条任意线的高程剖面图

对于许多肉眼看上去粗糙的样本而言，该方法是有吸引力的，如断裂和一些沉积涂层，但它并不适用于在许多应用中当前正产生的非常平坦的表面。典型抛光硅晶圆或精密加工过的镜面上的表面不规则性是典型的纳米级平面。

三种主要的方法被应用到这种表面。纵观以往，表面光度仪提供了一个工具来以趋近于纳米级

的精度精确地测量垂直高度。尽管得到了广泛使用，但表面光度仪对于许多表面应用而言有两个严重的缺点。其一是仅沿单个剖面求出高度。尽管这种纵剖面的分析相当直接，但它们对有着各向异性属性的复杂表面的适用性值得怀疑。以光栅方式扫描来收集点的阵列而形成一幅图像是可能的，但非常耗时。第二个限制是较大的针尖尺寸，限制了横向分辨率并使其不太可能精确地跟随陡坡或阶梯。

这使得干涉光学显微镜和 AFM 成为研究非常平坦表面的首选方法。前者是一个多世纪前所发现的光干涉原理的现代实现，而后者仅仅是近几十年发明并迅速商用的技术。

干涉光学显微镜（Robinson et al.，1991）将来自样本表面的光反射为经典干涉仪中的一条支路，然后与来自一条参考支路的光进行组合。图 1.55 显示了其示意图。相敏干涉的常规原理的出现使得路径长度的变化（由于表面上不同高度的点导致）产生了光的密度的变化。然后使用一台合适的 CCD 检测器阵列和模数转换器将该图像数字化，以便为表面上的每个点记录亮度值（及由其导出的高度值）。尽管使用的光的波长通常在 630nm 左右，但该波长的千分之一的干涉仪的两条支路间的相差会导致密度的变化，进而使得其垂直分辨率小于 1nm。但横向分辨率仍是 1μm 级左右，它受到所用光的波长及光学元件的设计的限制。

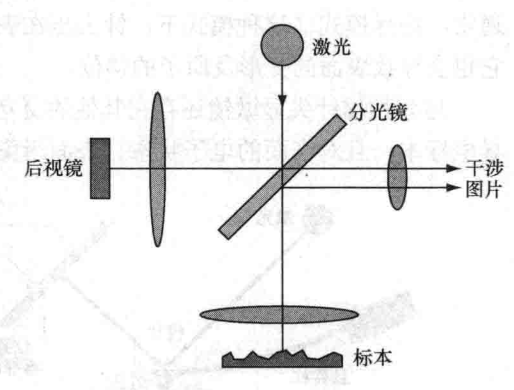


图 1.55　干涉显微镜示意图

因为没有光反射回检测器，如果表面有着高坡度或高镜面光洁度，则不能使用干涉测量光学显微镜。这样的点在最终的图像中会成为空缺点。填充这样的空缺点对于视觉目的可能是令人满意的（见第 4 章），但这当然会引起后续测量的偏差。

尽管横向分辨率远低于垂向分辨率，干涉显微镜仍然适合于许多应用。因为整个表面的对齐和该部分的形状或形式是未知的，隔得较远的点之间的绝对高度误差不能精确地测量。通常用一个函数来拟合数据的办法去处理对齐和形状的问题。对所有高程数据通过最小二乘拟合求出一个最佳拟合平面或其他低阶多项式函数，然后减去它，该过程称为“趋势分离”或“拆模”。于是就可以显示来自该表面的点的偏差的幅度。然而，隔得较远的点之间的绝对差受趋势分离面的影响，且区分那些隔得很开的点间小差别的能力被降低了。

图 1.56 显示了在商用干涉光学显微镜［翟柯（Zygo）公司］中用 Fizeau 和 Mirau 两组光学器件成像的一个非常平坦的表面（通过对硅晶圆抛光产生）。两幅图像的视域是相同的，高程的总垂直距离约为 2nm。图像底部的两个明亮的白点可能是由于光学器件中某处的污垢所引起的。这些人为假象与

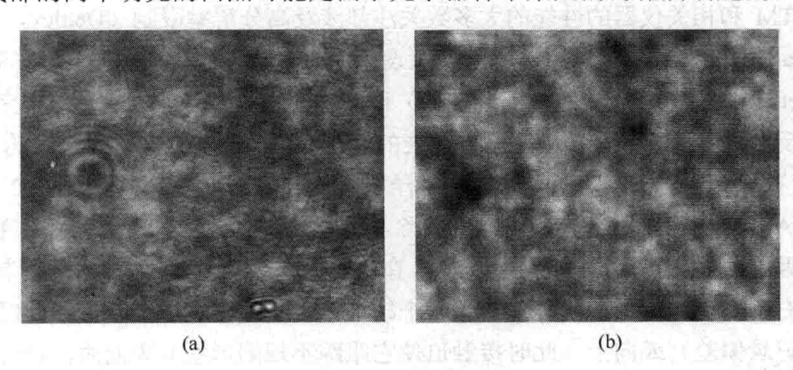


图 1.56　在平坦抛光的硅晶圆上使用 Fizeau 干涉仪(a)和 Mirau 干涉仪(b)的干涉距离图像的比较。由于扩散板与相机的检测器图案并未完全对齐，可以看到水平和垂直线的网格

Mirau 光学器件无关。此外，Fizeau 图像中围绕特性的振铃现象（振荡或波纹图样）在使用 Mirau 光学器件时，不再出现。这些特性通常可解释如下：Mirau 光学镜头在测量非常平坦的表面时更优。另一方面，Mirau 图像具有更低的横向分辨率（它看起来被“平滑”了）。

AFM 本质上是一台表面光度仪，它使用很小的针尖扫描样本表面的整个光栅（Wickramasinghe，1989）。标准表面光度仪的针尖因其微米级的尺寸，不能追踪表面上非常小或非常陡的不规则形状。AFM 针尖可更小、更锐利，尽管它通常仍不能精细到足以处理微电子电路和其他一些表面上出现的突变阶梯（或甚至底切）。针尖可以以针尖中或表面上围绕原子的电子之间交互的吸引模式或排斥模式操作。通常，排斥模式（这种模式下，针尖压在表面上）对于追踪较小的不规则形状效果可能要好一些，但它也会导致表面的变形及原子的移位。

这些扫描针尖显微镜还存在其他的交互模式。扫描隧道显微镜（STM）是原始的，但它仅可用于导电标本，且对表面的电子状态、表面污染和氧化非常敏感。随后出现的横向力、抽头模式和其他操作模式，可提供描述表面成分与属性等许多方面的能力（Birdi，2003），但原始的 AFM 模式最常用于确定表面的几何，理解起来相当复杂。

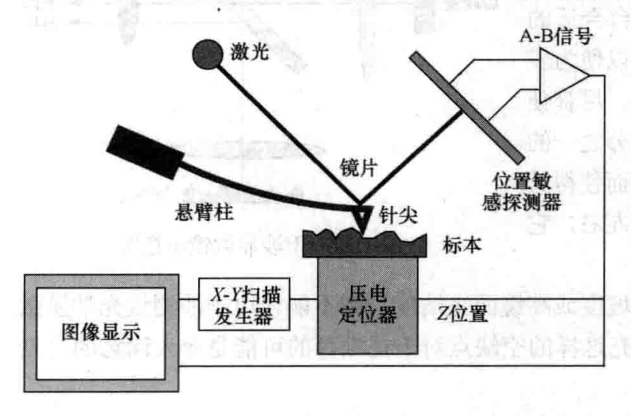


图 1.57　原子力显微镜（AFM）示意图

如图 1.57 的示意图所示，AFM 的经典操作模式是在 x、y 和 z 方向上移动样本。x 向、y 向扫描覆盖感兴趣区域，同时 z 向移动会使得笔尖回到由激光束反射确定的相同空位。在计算机中记录必要的运动来产生结果图像。改为移动针尖而不是样本，或使用线性阵列光传感器来测量针尖的偏转，而不是等待标本的 z 向移动，并不改变基本的原理。这一运动通常使用压电元件实现，其维数可通过改变施加的电压来灵敏地调整。但是，当在最大或最小尺寸（最低和最高频率）上丧失敏感性时，显微镜的调制传递函数（MTF）中会出现与这些运动相关的时滞。改用干涉测量而不是基于压电驱动电压来定位针尖的 x、y 坐标，可克服一些这类问题，但会增大设备的复杂性。

尽管 AFM 原理上有着几埃的横向分辨率，且在该范围有垂向敏感度，但并不总是能实现该性能。举个例子，为得到视觉上最好的图片而对扫描速度及放大器时间常数的调整，会消除一些精细尺度粗糙度，进而引起后续分析的偏差。相反，它会从电气或机械源引入额外的噪声。还需要特别小心来消除振动。

对 AFM、STM 和相关仪器的性能的大多数关注都涉及高分辨率限制（Denley，1990a；Denley，1990b；Grigg et al.，1992）。这通常由针尖的形状设置，而其并不容易描述。许多作者（Reiss et al.，1990；Pancorbo et al.，1991；Aguilar et al.，1992）已证明，采用与其他成像系统中完全相同的方法，通过对针尖形状反卷积并增强图像的清晰度是可能的。这类去卷积将在第 6 章中讨论。

图 1.58 显示了具有硬度压痕的抛光硅的一个样本（传统的粗糙度值的大小在 0.2～0.3nm 幅度，接近于干涉测量和 AFM 的标称垂向分辨率限制）。除灰度距离图像之外，该图显示了高度的等距渲染并应用了伪彩色，两者是这些仪器的常见显示模式。在压痕上，可看到 AFM 的一个限制，它表明压痕的左侧看起来要比右侧平滑。这种差别是扫描的一个缺陷，由于当跟踪一个向下（此时它可能滞后于实际的表面而不能记录偏差）或向上（此时接触迫使它跟踪不规则形状）表面时，针尖响应动力学是不同的。此外，压痕测量的深度要远小于实际深度，因为针尖不能跟踪压痕的最深部分，且 AFM 在垂直方向的校准精度要小于 x、y 方向的精度。

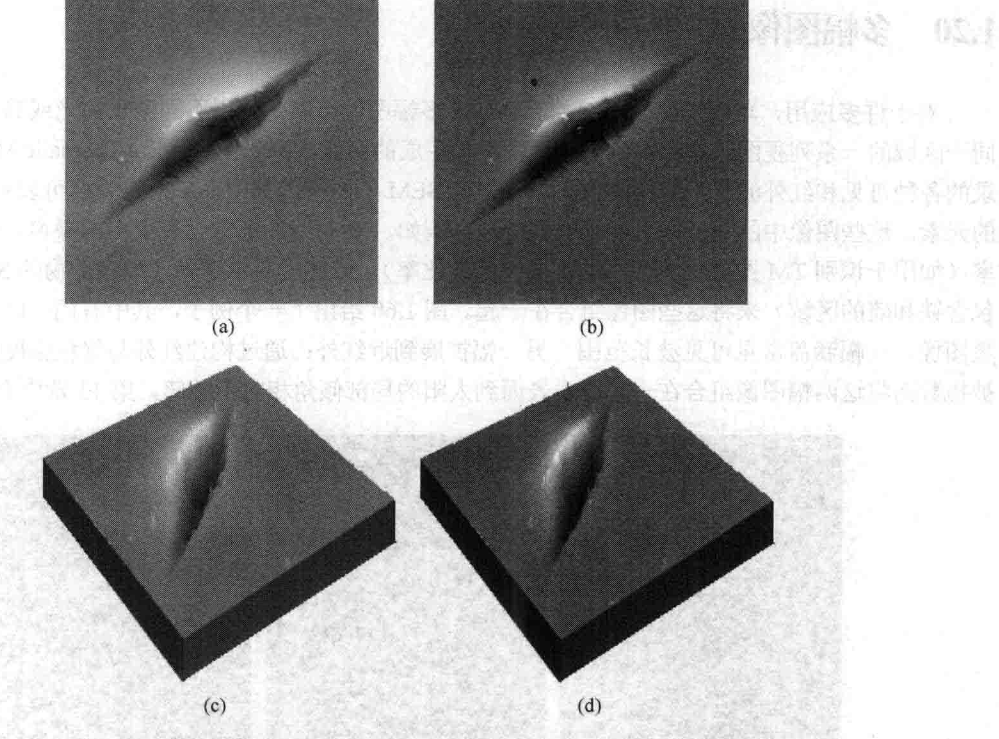


图 1.58　金属中努氏硬度压痕的 AFM 图像：(a)距离图像；(b)加入伪彩色的距离图像；(c)等距渲染图像；(d)加入伪彩的等距渲染图像

对于各种表面距离测量应用，存在着有足够的横向和垂向分辨率的描述表面的工具。干涉仪比 AFM 要更易于使用、以大范围的放大率工作，并能接受大的样本。由于设备特性，它引入更少的定向各向异性。另一方面，AFM 具有更高的横向分辨率，这对于有着非常精细特性的度量，正是所需要的，其在集成电路制造和纳米技术中，现在正变得很普通。

记住，许多仪器产生的信号仅与表面高程间接相关。有些情形下，它可表示样本的不同特征，如成分变化或电子属性。在图 1.59 中，表面高度的明显阶梯就是一个例子。该表面是平的，但样本的电子属性（微电子器件中的一个连接点）在信号中产生了变化。

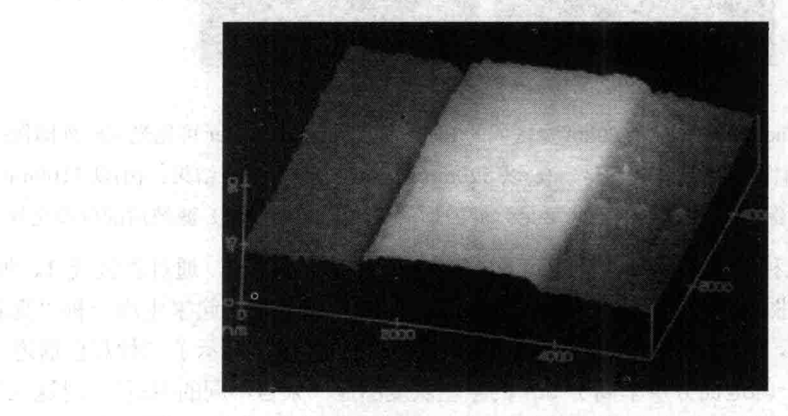


图 1.59　扫描式隧道显微镜（STM）图像。样本是表面平坦的硅，凸点是在 2μm 宽度的区域用植入的荧光粉改变的电子层

1.20 多幅图像

对于许多应用，单幅图像是不够的。例如，多幅图像可以是使用不同波长的光或其他信号而得到的同一区域的一系列视图。这样的例子包括由卫星生成的图像，例如由 Landsat Thematic Mapper（TM）记录的各种可见和红外波长生成的图像，以及来自 SEM 的图像，其中，X 射线强度可表示多达 12 个不同的元素。这些图像中的每一幅可能都需要处理；例如，X 射线图通常含有很多的噪声。然后使用各种比率（如用于识别 TM 图像中农作物的不同波长的比率），或使用布尔逻辑（如在矿物的 SEM 图像中定位包含铁和硫的区域）来将这些图像组合在一起。图 1.60 给出了一个例子，其中有同一区域的两幅彩色卫星图像，一幅涵盖常见可见波长范围，另一幅扩展到近红外；通过构建红外与绿色强度的比率来表示植被指数而将这两幅图像组合在一起，对表面到太阳的局部倾角相对不敏感。第 13 章中有更多的例子。

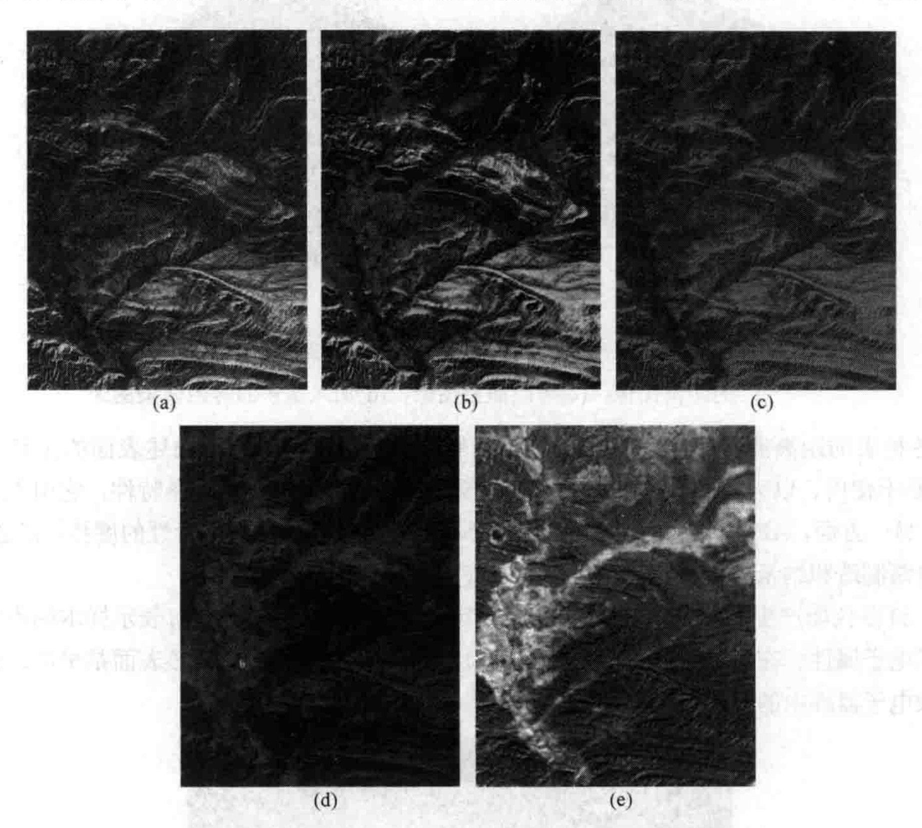


图 1.60　怀俄明州 Thermopolis 一个 20km 宽区域的 Landsat Thematic Mapper 可见光和红外图像：(a)可见光图像；(b)红外线图像；(c)以 520nm 过滤可见光图像的结果；(d)以 1100nm 过滤红外图像的结果；(e)绿色（可见）滤波后的强度与红外（IR）滤波后的强度之比

卫星图像通常是以几种波长获取的，涉及可见光和近红外波段。例如，通过将波段 1、波段 2 和波段 3 分别置入一幅彩色图像的蓝、绿和红通道，Landsat 图像可组合在一起来生成一种“真彩色”图像。尽管没有色彩敏感度，其他卫星会生成更高分辨率的图像。图 1.61 显示了三种彩色通道（其地面分辨率约为 15m），以及一幅地面分辨率高于 5m 的全色灰度图像（来自不同的卫星）。对这三个低分辨率通道组合将生成一幅彩色图像，将它转换为 HSI，然后将亮度通道替换为高分辨率图像，可生成如图 1.62 所示的结果，它同时具有色彩信息和较高的空间分辨率。彩色值有轻微移位且彩色边界不像亮度那样清晰的事实，并不会影响观看与解译。

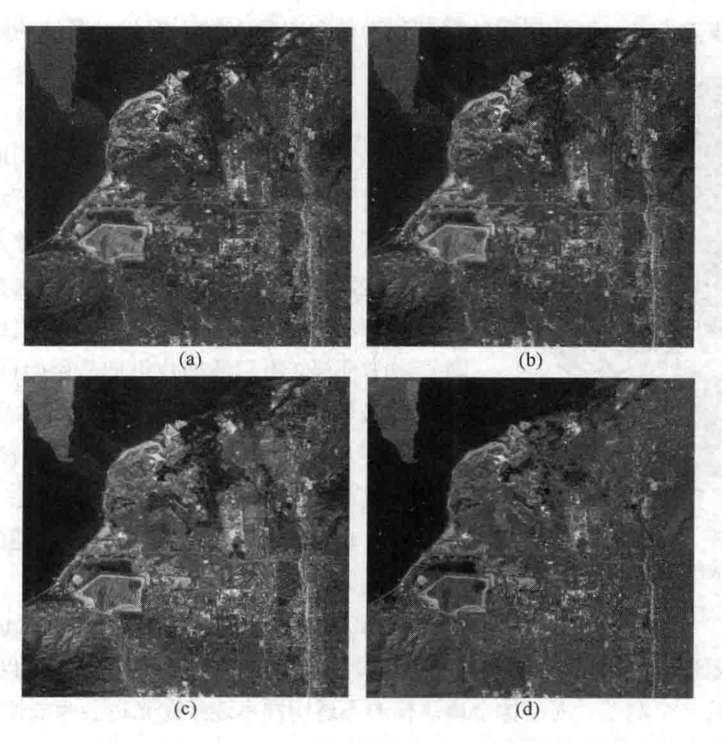


图 1.61　盐湖城的卫星图像：(a)波段 1（蓝色），15m 分辨率；(b)波段 2（绿色），15m 分辨率；(c)波段 3（红色），15m 分辨率；(d)全色，5m 分辨率

另一种多幅图像的情形是时间序列。这可能是用于跟踪天气系统运动的一系列卫星图像、追踪一颗彗星运动的望远镜图像、跟踪个体运动的监视照片，或一组用于跟踪细胞或纤毛跳动的显微图像。在所有这些情形下，所要求的通常是识别并定位每幅图像中的同一对象，尽管不同图像间的特征表现有一些逐渐变化。如果这些图像能被缩减到仅有少量特征位置的数据，则会大大降低对存储的需求。

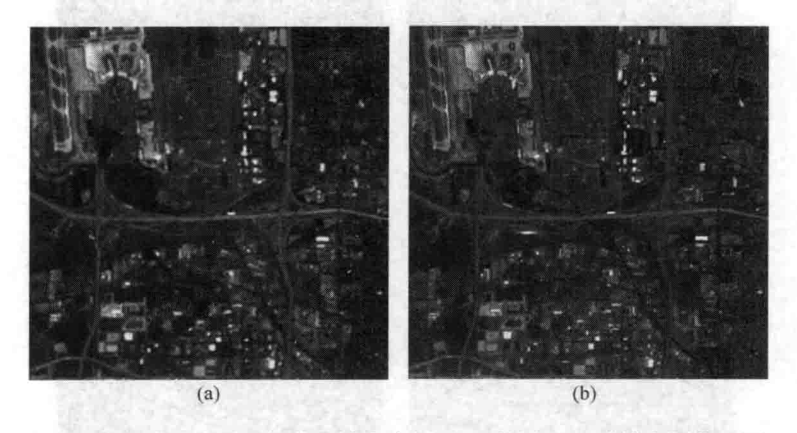


图 1.62　各个平面的组合（为显示细节已放大片断）：(a)使用 15m 分辨率图像，图 1.61(a)、(b)和(c)的 RGB 合成；(b)使用图 1.61(d)所示的更高分辨率的全色图像替换图(a)的强度通道

称为运动流的技术在一种较低的级别工作。使用这种方法，通过一种相关方法进行的像素模式的匹配，被用于创建显示连续图像间的运动的一个向量场。这种方法特别用于机器视觉和机器人作业中，这时连续图像之间的时间间隔非常小。简化向量场会得到适度的数据，这些数据通常被实时处理，故

存储不是问题。主要需求是，匹配所需的局部纹理信息必须存在。因此，图像的随机噪声应该很低，且每幅图像要包含有一致光照的简单且可区分的表面。匹配是在空间域或频率域使用互相关实现的（详见第 6 章）。

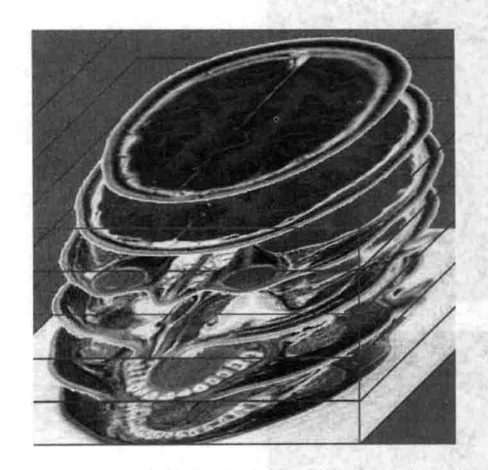


图 1.63　多个像素平面（通过人体头部的多个切片）组成的三维空间。体素是处理和测量三维图像的理想立方体

一组图像也可产生三维信息。它们通常是穿过一个固体的一系列平行的切片图像（见图 1.63）。如计算机断层术（CT）和磁共振成像（MRI）这样的医学成像方法就可产生这类数据。一些地震成像技术也如此。同样也在使用许多不同的连续截面方法的显微镜学中很常见。产生这样的一组图像的一种经典方法是，用超薄切片机把从原始样本得到的一组切片，分别在光学或电子显微镜对其成像，然后将所有图像对齐。

光学切片，尤其是对于 CSLM，有着很浅景深并能从透明标本深部收集图像，可消除对齐问题。除了可用于创建 3D 显示外，切片的对齐对于第 9 章中介绍的某些测量步骤也是很重要的。

像 SIMS（Secondary Ion Mass Spectrometer，二次离子质谱仪）这样的一些成像方法，通过对样本进行物理腐蚀，可产生一组深度图像，它也保持了对齐。图 1.64 显示了 SIMS 图像的一个例子。对于像金属这样的不透明样本连续抛光也会为成像产生新表面，但为了均匀地隔开它们而控制深度通常是困难的。聚焦离子束（FIB）加工（可在 SEM 内部完成，以使得标本位置保持固定并能被检查来控制该过程）可提供另一种方法来准备用于观看的多个表面。

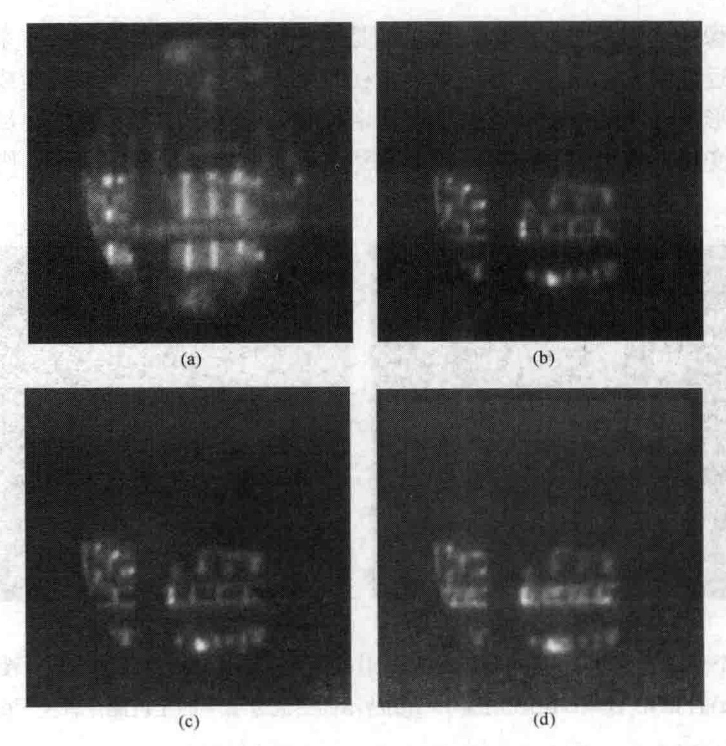


图 1.64　植入微电子器件的硼的 SIMS（二次离子质谱仪）图像。图像源自涵盖约 1μm 深度的 29 幅图像序列。每幅图像是通过物理地从表面上移去原子层得到的，该过程逐步侵蚀样品来揭示更深的结构

结构的三维解释的理想情形要求连续切片图像平面的横向分辨率等于平面间的间距。这会产生立方“体素”，它在三维上继承了方形像素在二维所具有的处理和测量的优点。但是，通常情形下，平面间的间距要远大于它们的横向分辨率，因此需要在平面间进行内插。在 SIMS 情形下，情况则与此相反，且平面间距（像一些原子尺寸一样小）远小于每个平面的横向分辨率（典型地为 1μm 左右）。

存在直接产生立方体体素图像的技术，如三维断层成像。此时，当样本旋转到不同的方向时，通常使用 X 射线或电子来获得一系列投影图像，然后数学重构计算每个体素的密度。得到的较大的三维图像阵列可被存储为一系列平面切片。在使用一个三维数据集时，存在许多的处理与显示模式，第 14 章对此进行了详细讨论。图 1.65 给出了个典型的显示，其中的剖面图像和表面重构都用于协助进行解译。

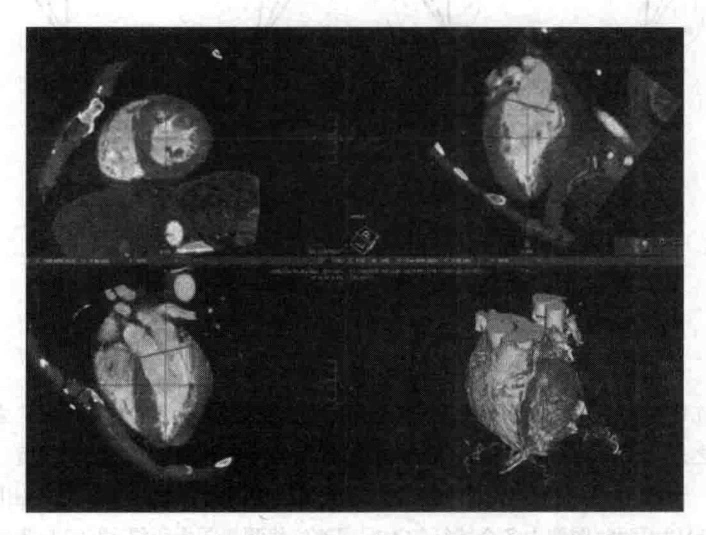


图 1.65　剖面与表面重构组合形成的医学图像

1.21　体视法

从相同场景的稍微不同视角拍摄的两幅图像中也可获得三维信息。人类体视法可助于对深度的感知，尽管还有一些其他的重要数据，如相对大小、先后次序、角度、大气霾雾和其他线索，文艺复兴时期的艺术家已非常有效地使用并理解了它们。和人类视觉的其他方面一样，体视法主要是比较性的：当注意力从一个特征移向另一个特征时，眼睛倾角的变化会向大脑表明哪一个更近。图 1.66 示意地显示了立体融合的基本原理，其中来自每个眼睛的图像被比较来定位每个视图中的同一特性。眼肌旋转眼睛，以使该特性落入中央凹，且肌肉向大脑提供倾向信息。注意，这表明体视法一次仅适用于视野中的一个特性，而不适用于整个场景。由所需倾向运动的数量，可确定对象的相对距离。由于仅进行比较性的测量，因此只需要方向或相对运动量来对每个对象的图像进行融合。

并非所有的动物均使用这种方法。例如，猫头鹰的眼睛在其眼窝中就不可移动。相反，其中央凹沿一条线有一个加长的形状，该线不是垂向的，而是与猫头鹰的脚之间有一定的角度。猫头鹰通过歪斜头部来完成融合（将感兴趣特性带到中央凹），并通过所需的歪斜来判断相对距离。

尽管人类视觉系统仅比较性地使用一个对象的图像的视差或倾斜，但在视野中测量两个对象的相对位置来计算它们的相对距离，或测量一个对象的倾角来计算它到观察者的距离，是有可能的。这通常在以从千米到微米的尺度范围来实现。

从两幅视图的距离信息来测量是三角法的一种直接应用。这两个视图中对象的横向位置是不同的，具体取决于它们的距离。从这些视差偏移来看，可通过称为体视法或摄影测量术的处理来计算。但是，图像的计算机融合是一项艰难的任务，它要求定位图像中的匹配点。试图基于局部纹理来匹配许多点的蛮力相关方法基本上类似于运动流方法。这会产生一些假匹配，但可通过后续的噪声滤波消除。另一种方法是定位图像中选取的“令人感兴趣的”点，这些点表示重要边界或特征边缘，进而更有把握地进行匹配。匹配点间的区域假设是平面或平滑样条。

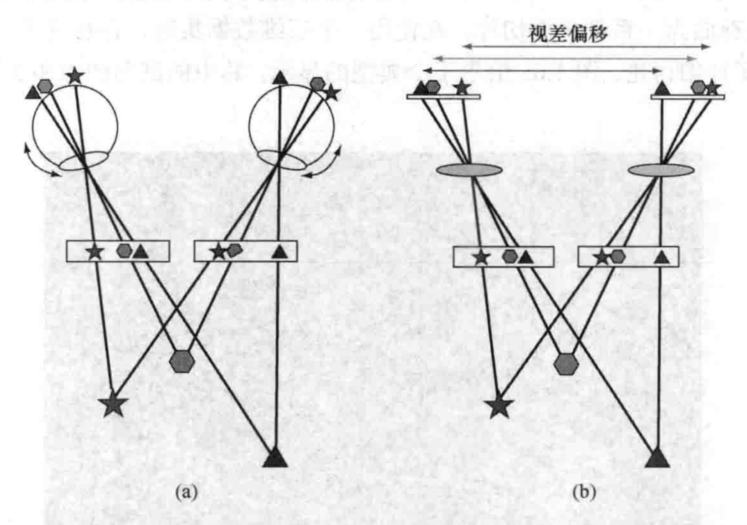


图 1.66　立体深度感知：(a)左眼视图和右眼视图识别的每个特征的相对距离，它由倾斜角之差给出，倾斜角是指眼球向内旋转使每个特征落到中央凹处的角度。每次完成一个特征。观看立体像对时，眼睛会得到相同的视觉线索并产生相同的解释；(b)使用两幅图像中各个特征的不同视差位移测量图像所得到的实际距离。必须知道两个视点之间的距离。在两幅图像中定位同一特征是自动分析任务最难的部分

不管融合是怎样完成的，两幅图像中点的偏移或视差给出了距离信息。这种方法用于表面高程制图，其范围可以大到用于产生地形图的卫星或航拍图片（此时，随着卫星或飞机位置的变化，两幅图像的拍摄时间相距一个很短的间隔），小到半导体芯片的扫描电子显微镜计量（通过倾斜样本，可产生两幅图像）。图 1.67 显示了 SEM 的一个立体对的例子；图 1.68 显示了由立体航拍相片创建的一幅地形图。

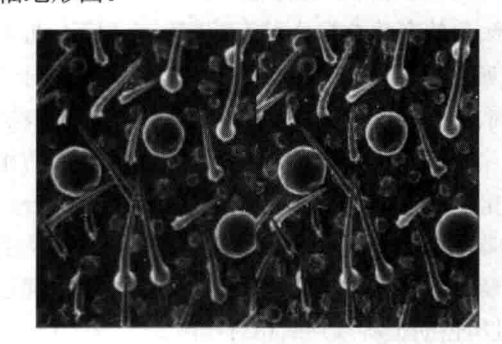


图 1.67　来自扫描电子显微镜（SEM）的立体像对。标本是一片树叶的表面；使入射到标本的光束倾斜 8° 后，由两个视角得到了这两幅图像



图 1.68　立体航拍图像上叠加的地形等高线图

用对世界的立体观看的能力来传达深度信息的结果，是使用立体相机来产生供观看的“立体感”幻灯片，这是 70 多年以前非常流行的行为，至今仍然在使用（见图 1.69）。立体电影（通常要求观看者戴上偏光眼镜）时常成为一种短暂的风尚，且现在已有某种兴趣推广家用立体电视显示屏。在出版物中用立体视图来说明科技论文现在已经相对普遍。一种格式将两幅图像相隔 7.5cm（人眼间的距离）并排放置，有经验的观众通过直视前方并让大脑来融合这两幅图像，不需要借助于光学器件就可看出立体效果来。另一种方法对每幅图像使用不同的颜色。以红和蓝（或红和青）套印同一幅图像，可使戴上彩色眼镜的观察者在每只眼睛中看到正确的图像，接下来人脑可分辨出深度信息（见图 1.70）。一些 SEM 显示器使用这种方法来显示表面的立体视图。

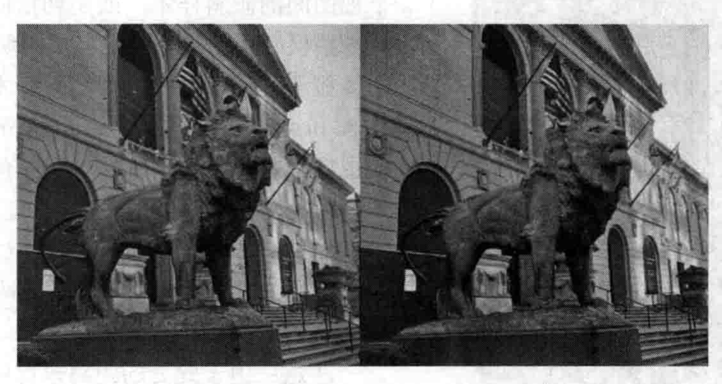


图 1.69　并排图像产生的立体视图（芝加哥艺术学院）

这种“立体”方法适用于灰度图像，但对于彩色图像（图 1.71 中所示），一个有趣的改动是通过将来自左眼图像的红色通道与来自右眼图像的绿色及蓝色通道进行合并，产生一种欺骗大脑看到全彩色立体场景的结果（对于主要使用左眼的观察者，该处理可能相反）。使用偏振光和镜片的图像投影也用于显示立体视图，但需要特殊的屏幕，因此对于在一个角度观看的人达不到满意的效果。现在正在推出呈现立体图像的各种新设备（Brown，2007）。

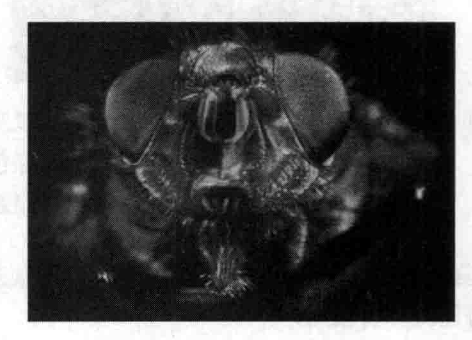


图 1.70　苍蝇头部的红色/青色立体图像。该图像是由两幅单独的 SEM 图像合成的。为观察图像，要在左眼前使用红色滤镜，而在右眼前面使用绿色、青色或蓝色滤镜

有几种不同的测量几何（Boyde，1973；Piazzesi，1973）。所有情形下，同一场景是从两个不同的位置或方向来观看的，并且这些位置间的距离或角度是精确已知的。有时，这可通过移动视点来实现，例如，携带相机的飞机。在航拍中，飞机的速度和方向是已知的，记录每幅图像的时间可给出飞行的距离。

在图 1.72 中，S 是位移距离（飞机飞行或相机移动的距离），WD 是工作距离或高度。出现在两幅不同图像中的两个点之间距离的视差 d_1-d_2（以平行于位移的方向测量）正比于这两点间的高度差。当 d_1 和 d_2 远小于 S 时，简化的关系为

$$h=\mathrm{WD}\cdot\frac{(d_1-d_2)}{S} \tag{1.10}$$

如果被测量的表面的垂直地势起伏 h 是 WD 的有效部分，则横向距离的透视收缩作为高度的函数也出现在图像中。图像中点的 x 和 y 坐标可用下式校正：

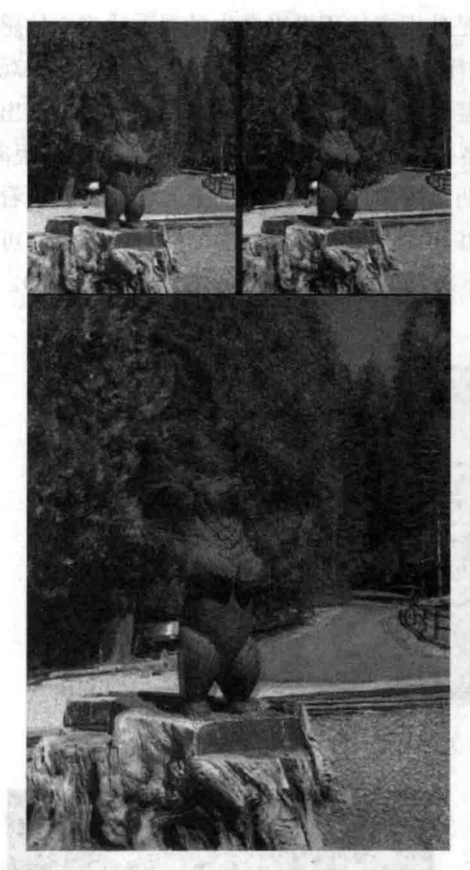


图 1.71　并排的彩色图像立体视图（上部），以及左眼图像的红色通道和右眼图像的绿色和蓝色通道的合成，可用红－青色眼镜观看

$$X' = X \cdot \frac{(\mathrm{WD} - h)}{\mathrm{WD}}, \quad Y' = Y \cdot \frac{(\mathrm{WD} - h)}{\mathrm{WD}} \tag{1.11}$$

如果两幅视图不在平行方向，而是朝内指向每个场景的同一中心点，那么两个出射点间的更大位移都可以达到。在航拍中，很少这样做，因为试图来对重叠图片进行拼接的大区域是不切实际的，且通常测量时不必获得足够的视差。但是，在较近的距离观察样本时（近景摄影测量或在 SEM 中查看时），通过在两个视图间的倾斜样本，或是朝指向场景中的相同中心特征拍摄两幅照片，就很容易实现了。在图 1.73 中，图像显示了通过沿垂直轴倾斜标本得到的两个视图。点 A 和点 B 由水平距离 d_1 或 d_2 分隔，而 d_1 或 d_2 在两幅图像中是不同的。从这一视差值和两幅图像间的已知倾角 δ，可计算出高度差 h 及连接这两点的直线（通常是由点定义的表面）的角度 θ（Boyde，1973），为

$$\theta = \arctan\left\{\frac{\cos\delta - d_2 / d_1}{\sin\delta}\right\}, \quad h_1 = \frac{d_1 \cdot \cos\delta - d_2}{\sin\delta} \tag{1.12}$$

注意：由于此处的距离以比例输入，所以角度 θ 不随放大率而变化。

当存在表面的同一区域的两个有角度视图时，可以使得特征的视差的相对偏移比它们的横向放大率大很多。这使得测量相对少量的表面起伏成为可能。5° ～ 10° 的角度最常用，对于非常平的表面，大到 20° 的倾斜角度可能有用。当对粗糙表面使用大的角度时，两幅图像都包含一些观察不到特征的阴影区，使得定位匹配点变得困难。并且，当一对图像的视差变得太大时，人类观察者在视觉上对这两张图片进行融合可能就变得很难。

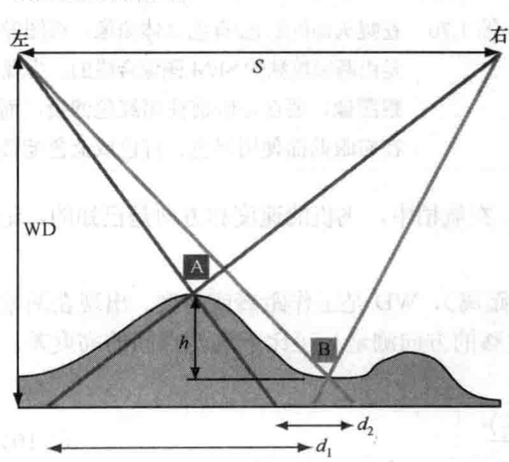


图 1.72　移动样本或视角得到的用于测量两幅不同图像中物体间垂直高度差的几何（通常用于航拍）

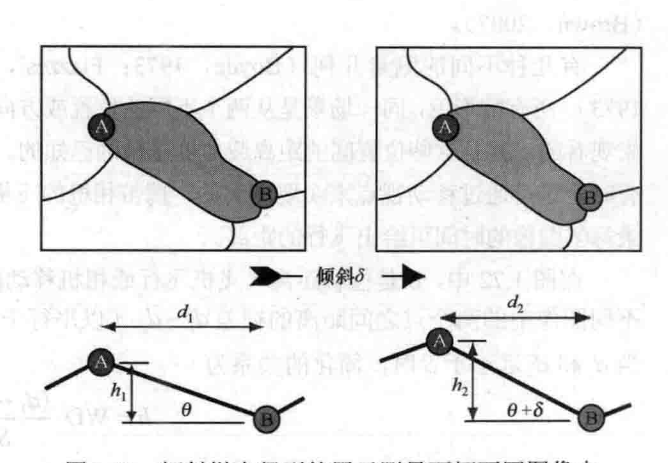


图 1.73　倾斜样本得到的用于测量两幅不同图像中物体间垂直高度差的几何（通常用于显微镜）

许多用 SEM 和航拍的立体对图像来完成的测量中用到了大量的人工交互。已经提出的自动融合算法需要用在后续章节中介绍的一些图像处理运算来使得识别每幅图像中的相同位置成为可能。当前，理解识别左边和右边图像的同一点，以及测量视差的原理足以给出高程。通过两幅图中许多不同点对的数据，就可以构造表面的一个完整图。可以用不同的格式来表现高度数据，比如将高度编码为阵列中每个像素的亮度或颜色的距离图像。其他显示模式，比如等高线、等角视图或阴影渲染，都可以生成来用以协助观察者解译这些图像。距离图像的数据以适用于多种类型图像处理的适当形式存在，并且适用于表面面积及表面上、下体积的测量。

当前制作的大多数地球表面的高程图是由来自航拍或卫星遥感成像图像的立体测量确定的（另一种广泛使用的技术是激光测距，该方法每次测量一个点）。当然，地球表面的 2/3 由水覆盖，因此不能用这种方法测量。海底部分通过侧扫描声呐立体测绘。这种技术与用于金星的雷达测绘技术很相似。不同之处在于，声呐使用声波（它能在水中传播），而雷达采用高频（毫米波长）电磁辐射，它能穿透覆盖金星的不透明的云层。

被麦哲伦探测器用于测绘金星的合成孔径雷达（SAR）并不直接给出高程数据来产生距离图像。SAR 的原理并不新［它是由 Carl Wiley 于 20 世纪 50 年代在固特异飞艇公司提出的］，也不仅局限于卫星和太空探测。例如，沙漠地区的航空测绘一直被用于穿透干燥的沙粒，并对底层的土地进行测绘来找到古河道。SAR 技术的原理是：移动的卫星（或其他平台）直接向下方并沿其移动轨迹的一侧发射出一系列短脉冲。平行于轨迹的方向被称为方位向（azimuth），垂直于轨迹的方向称为距离向（range）［注意不要和对高程信息编码的距离图像混淆］。“合成孔径”的名字表示，运动的天线有效地实现了一个更大天线的作用（在尺寸上等同于天线在脉冲传播过程中所移动的距离），这可以在方位向上更精确地分辨方向。

雷达记录返回脉冲的强度、传播时间和多普勒频移。强度是表面特性的测量，尽管雷达（或声呐）的反射率通常并不容易用表面结构来解释，而且并不与可见光的反照率或反射率直接相关。脉冲的传播时间给出了距离。对于一个非常平的表面，表面上存在一个到天线的距离相同的位置的弧度。由于卫星或飞机的运动，沿着弧度的每一点都会产生信号频率的不同的多普勒频移。测量每个返回脉冲的多普勒频移，可以得到弧度上的分辨率。地表的每一点以一个独特的距离和多普勒频移对返回脉冲起作用，从而可以重建出地面的距离图像。

然而，由于物体表面并不平，有多个位置和高程的组合可以产生相同的返回信号。将几次重叠扫描（对麦哲伦探测器而言，是几个连续轨道）的测量合并就可以优化高程数据。金星的数据库可以分辨地面相距 120m 的点，且分辨率在 120～300m 之间（取决于它们所在轨迹的位置以及距离轨迹中央的远近）的高程数据。对于每一点，其高程和反射率都被存储起来（见图 1.74）。对这些数据进行渲染时，高程数据被用来构造表面的形状，而反射率值则被用来给表面图像上色（见图 1.75）。当然，这些颜色并不是其表面的实际视觉外观。

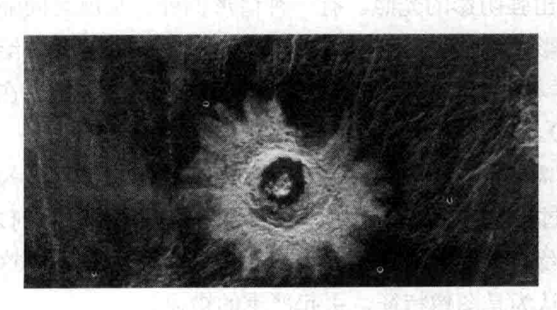


图 1.74　合成孔径雷达得到的金星距离图像



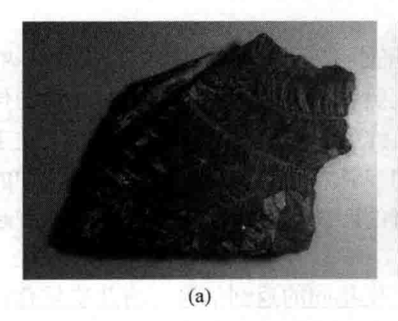
图 1.75　由 SAR 数据渲染的金星表面

使用雷达和声呐进行合成孔径测距并不是这些信号被用来制作距离图的唯一方式。通过垂直向下发射电波并测量反射波或返回时间，给出了某一单个点的距离。许多这样的测量都可以用来制作地图或图像。如果测量船以光栅模式来回移动，从而覆盖整个湖泊表面，可用简易的“鱼探仪”型的声呐系统来绘制湖底表面图像。也可以用其他信号。Clementine 探月计划用一种类似的方式采用来自轨道太空飞船的激光束。这种激光测距也被用于地球的航空测绘中。通过以脉冲方式发射激光并等待回波，Clementine 号测量了整个月球表面间相隔几百米的一些位置，其垂向分辨率约为 40m。

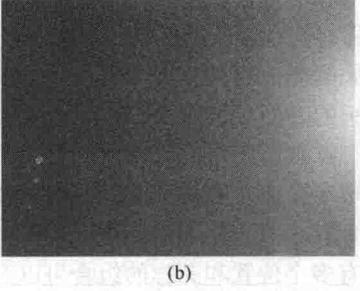
1.22 成像需求

鉴于上面描述的图像类型和图像源的多样性，有一些常用准则可用于为了进行计算机处理和分析目的的图像。首先需要全局一致性。无论在图像中的哪个地方，同一类型的物体或结构看上去应该是一样的。这就意味着亮度和色度值应该保持一致，因此，对不同时间获得的图像而言，其照度必须一致且稳定。当表面不是平面时，例如从卫星上观察到的地球表面或显微镜观察到的断裂面，可能有改变局部方向的校正，但是通常需要计算，或需要表面和光照源的先验知识。

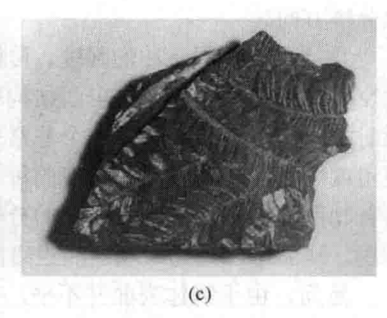
图 1.76 给出了非均匀光照图像的例子。拍摄相同光照下没有样本存在的翻拍台的底部，得到“背景”的图像，从而使得这种不均匀性调得均匀。从源图像减去或除以（取决于相机具有对数还是线性响应）这个背景图像。这类调均匀将在第 4 章关于校正图像缺陷的内容中介绍，同时在不能直接获取它时，还会给出估计背景图像的其他方法。



(a)



(b)



(c)

图 1.76 (a)非均匀光照下得到的宏观图像；(b)从翻拍台上移去样本后，在相同光照下记录的背景图像；(c)原始图像像素除以背景图像像素，并扩展结果图像的对比度后所得到的调匀图像

这种对均匀的需求，限制了通常成像的表面类型。平面或至少简单且已知的表面，比复杂的表面更容易处理。比起有着拱门、桥梁、洞穴和圆环等遮蔽了某些结构的平面而言，简单连接的平面更容易解译。具有位次问题的特征（其中的某些特征完全或部分地隐蔽在其他特征后面）给解译和测量带来了困难。在大多数情况下，也不希望出现投射出强阴影的光照。有一种情形例外，那就是间隔合适的特征投射的阴影之间不会相互干扰。利用物体影长和照明几何可以计算特征的高度。图 1.77 给出了航拍中的一个例子。为 TEM 准备的一种形式的样本从点源沉积了一个金属或碳薄膜，该点源也在颗粒或突起后面产生可以同样利用的阴影（见图 1.78）。

除了全局一致性，对变化的局部灵敏性也是需要的。这意味着，边缘和边界必须被很好地勾勒和精确定位。相机传感器的分辨率在前面已经讨论过了。总体而言，在信号链中会削弱高频的任何东西，都会影响随后用于测量的辨识特征边界或定位边缘的能力。另一方面，比如光学器件上的灰尘这类简单问题，都会引入局部变化，而该变化可能被误认为是图像特征，引起严重的错误。



图 1.77　航拍照片，由阴影的长度和已知太阳位置可计算出树木的高度及木材堆置场中原木堆的高度，进而估算出木材的数量

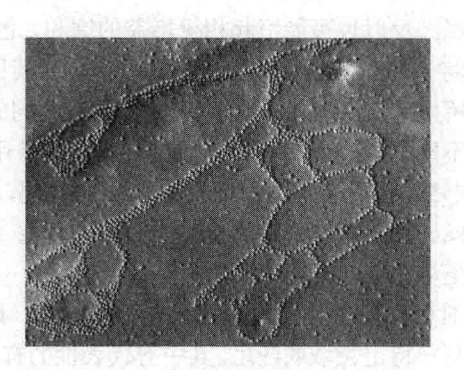


图 1.78　组织冷冻断裂的电子显微镜图像，显示了细胞间间隙的轮廓

尺寸测量需要清晰了解成像系统的几何。知道显微镜的放大倍率、相机镜头的焦距或卫星的高度通常是直接的。可以利用光学器件的知识，也可以使用已知尺度的图像或标准图像完成横向尺度校正。当观看几何更复杂时，或者因为表面不是平面，或视角不是垂直向下时，测量就更加困难，且需要首先确定其几何，或者确定被观测物体上的规模或基准标记的内含物。

图 1.79 显示了最简单的一类失真，它在从一个角度观看平面时出现。图像的不同部分有不同的放大尺度，使得后续分析变得更加困难。失真也使得将复杂平面的多幅图片组成为一张拼接图变得复杂。该问题出现在不同尺度的应用中。只需利用图像扭曲使得边缘配准，星球表面的卫星图像就可组合成覆盖大面积的拼接图。这类扭曲将会在第 4 章中讨论。因为整个表面几何是未知的，粗糙表面的图像用这种方式进行组合就更难了。

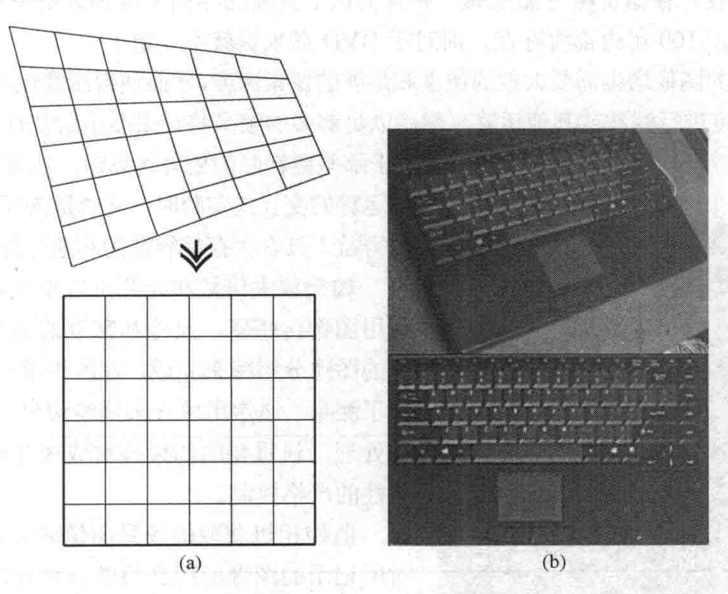


图 1.79　从远离表面法线的位置观看表面时，所出现的几何失真。校正该失真得到了一幅矩形图像，矩形图像可正确地处理和测量，或与相邻图像拼贴在一起，这样做需要知道可用来确定矩形的观察几何，或在场景中包含了已知的基准标记

测量亮度信息（如密度和色度值）时，需要一个稳定的光源和传感器。所记录的颜色受到白炽灯泡色温变化的影响，该变化可能是由于电压小幅波动、灯泡升温或老化引起的。由于荧光灯管高频闪

烁，荧光灯照明，尤其是和 X 光片或密度凝胶一起用于灯箱中时，可能不稳定或在固态相机中引入干涉。亮镜面反射也可能引起相机增益的饱和、散射或偏移。

当将一幅图片数字化到计算机时，谨记其目的将会给你带来帮助。下面列出了可能的几种情况，它们对所用到的软/硬件的限制和要求也是不同的。后续章节会对这些内容进行详细分析。贯穿本书的重点是不同处理和测量技术产生的结果，以及用简明的语言对方法的描述和比较不同方法的说明。有许多的书籍给出了实现不同算法的计算机代码，比如 Russ & Russ，2008；Parker，1997；Ritter & Wilson，2001；Pavlidis，1982；Myler & Weeks，1993；Umbaugh，1998；Seul et al.，2000；Lichtenbelt et al.，1998，更多书籍定期会出版。

同时也有大量数字图像处理方面的书籍，其中一些书籍深入研究这门科学的数学基础，另一些书籍注重某一特定领域的应用。其中有代表性的有 Gonzalez & Woods，1993；Pratt，1991；Rosenfeld & Kak，1982；Weeks，1996；Pitas，2000；Nikolaidis & Pitas，2001；Costa & Cesar，2001；Sonka et al.，1999；Kriete，1992；Hader，1992；Sanchez & Canton，1999。

一些以应用为主的书籍介绍了图像处理以及较少篇幅的图像测量在所研究特定领域的应用。几个这样的例子包括取证学（Russ，2002）、食品科学（Russ，2004）、地质学（Francus，2004）和医学成像（Costaridou，2004；Rangayyan，2005）。另一些书籍采用特定的软件程序，如 Photoshop（Russ，2009）和 MATLAB（Gonzalez et al.，2004）来介绍处理和测量。本书的重点是这些方法的一般性，表明同样的技术应用到广泛的图像类型中，并试图教会读者通过性能和结果而不是理论来选择方法。

由于大容量存储设备（可读写 DVD 盘）价格的下降，以及不同地点需要图像的多个档案副本，使得图像的存储和归档技术变得越来越具有吸引力。在许多情况中，该应用还包括存储图像的硬拷贝印刷以及将图像传送到其他位置。如果不需要对图像进行进一步的处理和测量，那么图像压缩是可接受的。电子存储的优点是，图像不会随着时间的推移而退化，并且通过适当的归档和索引程序可进行存取。另一方面，胶片存储价格更加低廉，并且提供了更高的存储密度和分辨率，更为可能的是，用来观看胶片的设备在 100 年内都将存在，而对于 DVD 盘来说就不一定了。

用于视觉目的的图像增强需要大量的像素和足够的像素深度，从而使得图像包含足够的信息而获取，进而以较高的保真度进行滤波或其他运算，然后以足够多的细节将结果显示给观众。光照的均匀性和几何的控制是次要的。当使用大图片时，尤其是对于那些更耗时的处理运算时，或是需要进行带有许多不同运算的交互式实验时，又或是需要实时结果时，该应用就得益于具有大存储容量的高速计算机或专用硬件。



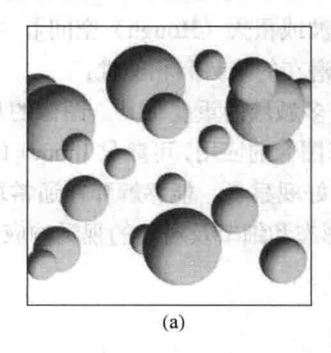
图 1.80　反投影摄影测量。将标准器放在场景中物体之前所在的位置，使用相同的相机得到其图像，并将该图像叠加到原始图像上，就可测量出物体的高度

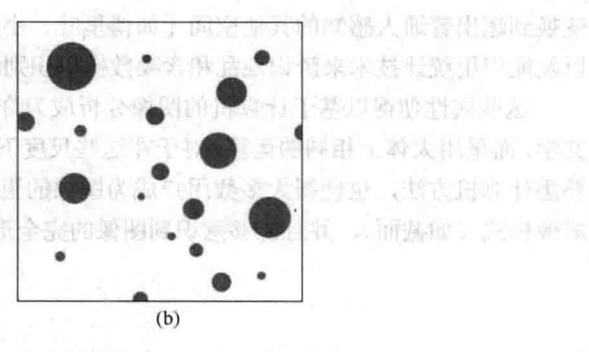
如果放大倍数和光照可以事先校正和调整来最佳地利用图像传感器，尺寸和密度值的测量通常就可以用适中的图像分辨率来完成。在测量前可能需要处理（比如，为了测量，通常用求导来描绘边界），这通常完全在软件中处理。最重要的约束是对成像几何及光照均匀性和持续性的严格控制。

监控相机获取的场景图像通常需要几何测量。尽管对所记录的图像应用必要的三角计算是可能的，但采用反投影方法通常会更好。如图 1.80 所示，该过程事后将校准标准放置在同一场景中，通过图像的叠加来进行直接比较。该技术在取证应用中特别有用，因为它非常直观，不需要复杂的解释说明，也不需要对透视收缩进行校正或其他假定。

质量控制应用不涉及绝对测量，甚至检测变化。在许多情况下，这通过将参考图像从每个获取的图像中逐点地减去，来检测总变化。这可以通过模拟电路实时完成。需要解决的主要问题是，如何防止相机或目标位置的意外变动以及光照引起的变化。定位诸如孔洞和成分通常用互相关来实现，并由专用处理器来应用。

二维或三维空间的微观结构研究通常从图像测量开始，具有和上面一样的需求，此外还需要使测量值服从于适当的体视学和统计分析的能力。对于平面截面或投影图像（见图 1.81）来说，按结构来解释图像是不同的。后者接近于人类视觉，而前者则不然。如图 1.82 所示的截面图像在三维中包含用于测量的信息，它们可以通过适当的分析揭示。第 9 章描述了该方法。如图 1.83 所示的投影图像，其视觉上更熟悉，但是可能不容易进行解释。绝大多数医学成像技术对应于以上两种方法中的一种（CT 和 MRI 图像是截面，传统的 X 射线是投影）。





(a) (b)

图 1.81 (a)人们熟悉的球体投影响图像显示了特征的外表面，但有些特征则被部分遮盖或被完全遮盖，因此很难求出其数量或尺寸分布；(b)人们不太熟悉的截面图像并未最大限度地显示出各个特征，但统计上可以预测球体的尺寸分布和数量

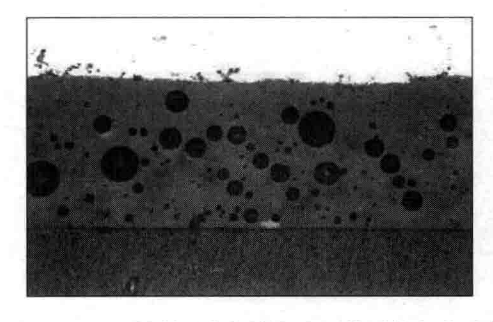


图 1.82 钢材所用彩色搪瓷涂层截面的光学显微图像。球形气泡是在搪瓷层烧制过程中出现的。截面表明，圆的直径小于球体的最大直径，但由于气泡的形状已知，因此可推出球体的数量和尺寸分布

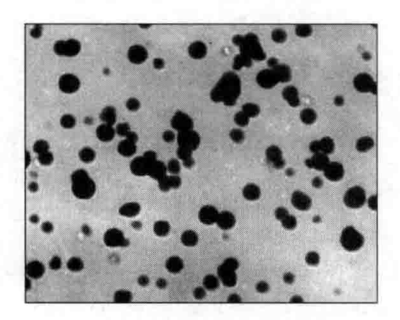


图 1.83 透明厚截面中的乳胶球体的透射电子显微镜图像。有些球体被其他球体部分遮盖。若已知截面厚度，则可算出尺寸分布和球体所占的体积比，但无法算出完全被遮盖的一些小特征的尺寸分布和体积比

三维成像使用大数据集，而且在许多情况下，二维图像的对齐是非常重要的。一些三维结构参数能从二维图像中推导出来。另外，主要的拓扑信息只能从三维数据集中得到。三维的处理和测量运算对计算机容量和速度的要求很高。以人类可以解释（如果不总是熟悉）的方式显示三维信息的技术在不断改进中，但仍需要进一步发展。它们也对处理器和显示速度的要求很高。第 13 章和第 14 章讨论了三维成像。

分类和识别通常被认为是基于计算机的图像分析的高端任务。从复杂性来说，其范围包括定位和

识别属于几个确定类别的孤立物体，到更多开放性问题。前者的例子有，用于机器人操纵的目标定位或在监控图像中识别目标。后者的一个例子是医学诊断，其中重要的信息主要来自于信息源而不是图像本身。模糊逻辑、专家系统和神经网络都在一定程度上成功地应用到这些问题中。这些方法将在第 12 章中介绍。将正确的信息从图像中提取出来并送入决策逻辑，比简单的处理和测量更加困难，因为对于特定应用的最佳算法，其自身必须作为整个过程的一部分来确定。

所有这些任务都需要基于计算机的图像处理和分析系统，并通过获取图像的硬件设备，来仿真人类的视觉系统。在许多情况下，它们以复制被认为用于视觉中的算法来这样做，但在其他情况下，采用了完全不同的方法。尽管没有基于计算机的图像系统在灵活性上或速度上能接近复制人类视觉的整个性能，在一些特定的任务中，基于计算机的方法胜于人类。例如，它能检测更多的成像信号而不仅仅是可见光；它不受外界的影响以及疲劳或分神的影响；能完成绝对测量而不是相对比较；能将图像变换到超出普通人感知的其他空间［如傅里叶、小波或霍夫（Hough）空间］，并从中提取隐含信息；以及能应用统计技术来辨识混乱和含噪数据以识别潜在的趋势和相似性。

这些属性使得以基于计算机的图像分析成为许多领域的重要工具。图像的尺寸可能是从微观到天文学，而采用大体上相同的运算。对于在这些尺度下图像的应用，可参见 Inoue（1986）和 Sabins（1987）。熟悉计算机方法，也使得大多数用户成为图像的更好观察者，能够解释在通常场景中很少见的非寻常成像模式（如截面），并且能够意识到图像的完全形态和细节及对其的视觉响应。

第2章 人类视觉[1]

在一本主要讨论数字图像的计算机处理和测量的书中，用一整章篇幅来讨论人类视觉特性有两个主要原因。首先，许多图像处理涉及提高图像中细节及特征的可见度，而这取决于人们对从图像中所看到（及所忽略）的内容的一些理解。其次，在后续章节描述的用于图像处理和场景中目标检测的许多算法，或多或少都基于我们对人类视觉处理如何工作的理解。当然，这并不是处理算法的唯一源头。有些是基于与样本相互作用的光（或其他信号）的物理性质，有些只不过是被发现有用的特别程序。但是，从视觉研究取得的理解，也已成为计算机处理方法的一个重要源泉。

本章还旨在拓宽读者理解人眼与数码相机的硬件之间，以及计算机与人类神经系统的软件之间的异同。这些知识有助于我们成为更细心的观察者以及这些新工具的更好用户。

2.1 我们看到了什么，为什么会看到

人类就是视觉动物。人类获取的大部分信息是通过眼睛（及大脑中的相关回路）而不是通过触觉、嗅觉、听觉或味觉实现的。不管是好还是坏，这也是科学家通过实验获得信息的途径。但是通过几百万年的进化所生成的解译图像的技巧在处理科学图像时就和“真实世界”所感受的有所不同。了解要提取的信息类型的差异、期望的作用和熟悉度，以及由我们的视觉系统引入的偏差，对于相信其结果的科学家来说，是一个必不可少的要求。

对于感官机能未受任何损害的普通人而言，通过视觉途径流过的信息占 90%～95%。事实上，我们对视力的依赖程度，可以通过对那些视力不理想或随年龄增长视力恶化的人的校正手段（从眼镜到激光眼科手术）的可行性来判断。对于重度听力受损者，助听器与人工耳蜗植入是存在的（但使用得不多），但对于其他感官则没有治标的方法。当味觉变得不那么敏感时，唯一的解决办法是洒多一些辣椒粉。

并非所有的动物，甚至所有的哺乳动物，都达到人类依赖或使用视力的程度（见图 2.1）。蝙蝠和海豚利用回声定位或声呐来探测它们周围的世界。响尾蛇感应红外辐射。生活在地下的鼹鼠用鼻子周围的敏感触觉器官代替视力。猎犬跟踪气味，蝴蝶具有非常敏感的味觉器官，甚至可以察觉到单个分子。有些鳗鱼产生并感应电场来与周围环境互动。鱼和短吻鳄具有压感，能在其周围含水的环境中感知非常轻微的运动。鸟儿和蜜蜂都有能力来检测光的偏振，作为在阴天时定位太阳的一种辅助。鸟类和一些细菌能够感知地球磁场的方向，来作为导航的另一种辅助。而许多鸟类和昆虫具有能检测超出我们视觉范围之外的红外或紫外光的视觉系统（Goldsmith，2006）。

对人类而言，很难想象在蝙蝠、鳗鱼或鼹鼠眼里，世界是什么样子，即使单词“想象”说明了这一问题。“图像”表示在大脑里面构造的一幅图画或场景。图像是能被我们中的大多数理解的唯一的世界数据的组织形式，正如我们对图像的依赖一样，我们的语言反映或说明了这种偏差。与天生就失明的人士访谈表明，他们通过触摸、声音和其他线索建立起一种精神意象。

---

① 本章的部分内容最初以一系列的文章出现在 2004 年的皇家显微镜学会会议论文集（*Proceedings of the Royal Microscopy Society*）中。

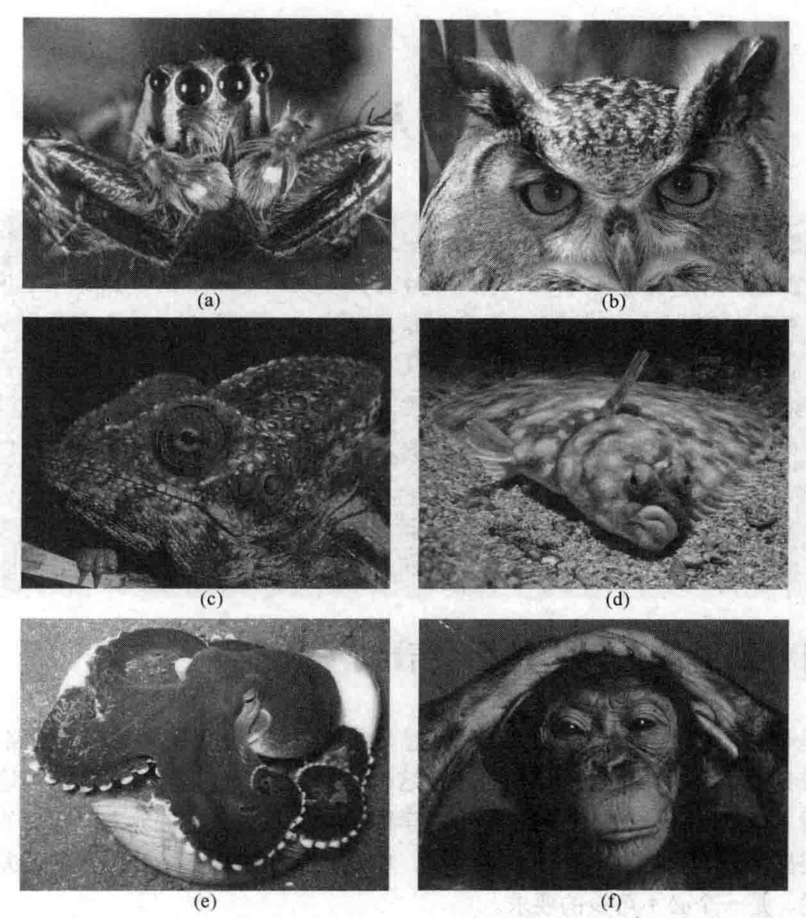


图 2.1　眼睛有许多形式，并针对不同的目的进行了优化。蜘蛛的眼睛由具多组晶体和传感器构成，其分辨率较低，但覆盖范围较大。鸟的眼睛有较高的敏度和分辨率。变色龙可单独旋转其眼睛来跟踪左右视野内中的不同对象。成年比目鱼在头的一侧有两只眼睛。章鱼的眼睛有很好的颜色灵敏度，但神经网络朝与哺乳动物视网膜相反的一面进化。灵长类动物也非常适应于立体视觉，且与其他大多数哺乳动物相比，对红色有更好的灵敏度

两个向前直视的眼睛能够检测波长范围 400～700nm（从蓝色到红色）的光。人类起源于树栖灵长类动物，这种动物依赖于导航和狩猎的视觉和体视。相反，许多动物和昆虫牺牲体视而增加覆盖范围，即其双眼相距较宽来检测运动。少数动物如变色龙，可以独立移动它们的眼睛来跟踪不同的目标。但即使对具有体视的猎人而言，也有许多鸟类，其视力远优于人类。鹰具有可以辨别近一英里外的一只老鼠的分辨能力。事实上，大多数鸟类的眼睛在头部占有的比例比人大。一些鸟类的眼睛是如此之大，以至于影响其他功能，如需要眨动眼皮，迫使眼睛下降到喉咙才能够吞咽食物。

经常被引用的谚语“百闻不如一见”，被用做图像和丰富信息内容的重要性的说明。但在许多方面，这种说法是错误的。首先，在计算机中，数字化并存储的一幅典型的图像占据了几百万个单词的文本空间。用本章作为一个例子，不包括格式信息，1000 字需要 6080 字节来存储，并可以在不丢失任何信息的情况下被压缩到 2150 字节/千字。此外，现代数码相机的分辨率远远不及人眼，人眼拥有约 1.6 亿个视杆细胞和视锥细胞。另一方面，作为从一个人到另一个人信息交流的方式，图像是非常低效的。在没有一些支持信息将图像引入他们的视野并创建上下文来解释它时，期望另一个人从图片中获得和

我们相同的信息是不合理的。Arlo Guthrie[②]在歌曲 *Alice's Restaurant* 中将其描述为“27 个 8×10 的有圆圈和箭头的彩色光泽图片，并且每一张背面都一段话。”这并非许多典型科研论文中的较坏描述！

研究表明，我们看到的图像受到了文化差异的强烈影响。许多西方人将他们的注意力固定在一个（或几个）前面的或颜色明亮的物体上，而忽略周围环境。许多亚洲人更注重整体的场景和背景细节，注意到前面存在的物体，但并没有将特别的注意力放在学习或记住其特征上。当然，在一定的场景中识别的东西，如果对观察者熟悉的话，将会严重影响到注意力集中的位置。

人类视觉，可以从图像中提取到几种不同类型的信息，而且发生的很多处理已通过演化和经验优化来有效地执行。但同时，其他类型的信息被忽略或抑制，不能正常观察。福尔摩斯经常批评华生“看但不是观察”，而这正是任何光子落在视网膜上与触发意识心智的一个很好的区别。本章探讨一些信息提取和心灵意识水平预警的过程。提取过程忽略了某些种类的信息，或使某些场景内容非常难以察觉。而期望影响感知——我们更容易发现正在找的东西，反之亦然。

2.2 识别

人类视觉的不少目的是识别。无论是寻找食物、躲避捕食者，还是引起配偶的注意，在一幅图像中首先引起我们注意的是那些熟悉的物体。必须认识到，一个物体或特征必须有一个名字——一些我们意识可以分配的标签。在这个标签的背后，是物体的心智模型，物体可能表示为文字、图像、和事物相关的记忆或其他形式。该模型捕捉到物体（对我们）重要的特性。遗憾的是，在许多科学实验中，分配给人类视觉的任务很少是识别熟悉的物体，而是检测和描述陌生的物体。

不存在标签的事物是难以识别或辨认的。相伴的存储表示是包括了那些在我们每个个体记忆中的特性或特征的例子或实例，这些特性或特征是这类物体的识别标志。当然，每个人都可能有自己的例子。引用一个简单的例子，一个人认为的“海豚”可能是哺乳动物（“鳍”），而另一个人则可能认为是鱼（鲯鳅）。错误地指定了本意导致误解。此外，对应于标签所记住的例子通常并不代表任何真正的对象，也不是真实对象的任何一种统计上的有代表性的组合，而是包含那些通过以往的经历事件对每个个体重要的特性和特殊属性。换言之，即使我们对物体的标签下降到相同的属、种及品种，它们也可能并不由相同一组期望的特征来支持。当科学家发表用图来进行说明的论文时，图代表科学家对这类物体的个体模型。既然“真实的”图像被用到，由于它们显示了可能在任何意义上不可能被表示的类别的特殊成员，就说明可能不具有太大的代表性，既不对应于作者的精神表示，也不对应于这类原型的统计意义。

人类视觉根源的基本技术是比较。图像中没有东西是通过眼睛和思维测量的；在我们的大脑中没有尺子和量角器来测量大小，没有分光光度计来测量颜色和亮度。具有相似方向、环境和光照并可以彼此相邻观看的特征比较起来最容易。那些必须在大脑中翻转或旋转的特征则相对困难一些。图 2.2 显示了一个例子，其在大脑中将每个目标翻转来匹配对齐并确定哪些特征相同，以及哪个是镜像图像所需的时间长度正比于两者之间的角度差。和记忆相比，

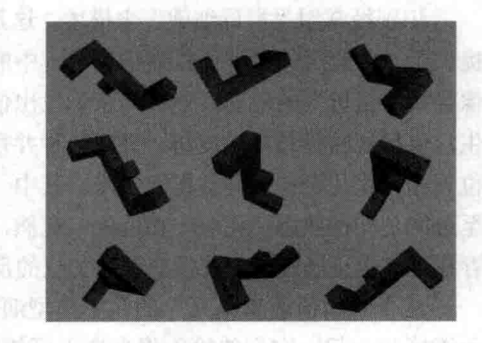


图 2.2 这些物体中的有些是相同的，有些是镜像图像。在大脑中旋转每个物体进行比较所需的时间长度正比于角度差

② 美国民谣歌手。——译者注

以同样的方式工作需要时间。如果所记住的物体是非常熟悉的，那么潜在的模型由可以比较的特征列表组成，这会非常快。毕竟，这就是识别如何工作。

如果记住的对象不熟悉，也没有标签和模型，那么比较只是取决于原始视图中哪些特征被记住了。记忆过程进行得如何好，哪个特征和特点被选来回忆，仍然受限于同其他模型的比较。举个例子，目击证人对犯罪和事故现场的描述十分不可靠。基于记忆中来自其他物体的相似性或差异，不同的观察者认为场景中或嫌疑犯身上不同的属性是引人注意的，因此每个人的结果是不同的。警方对嫌疑犯的画像很少能与其被抓获后所拍摄的实际照片很好地匹配。在某些方面，它们是漫画，所强调的是对单独观看者要么熟悉要么不寻常的（通常是微不足道的）某些方面（见图 2.3）。对一个可能的银行抢劫犯的信息是显而易见的：提供一些事后可能会被丢弃的、容易记住的线索，如文身、假发、绷带和跛行等。

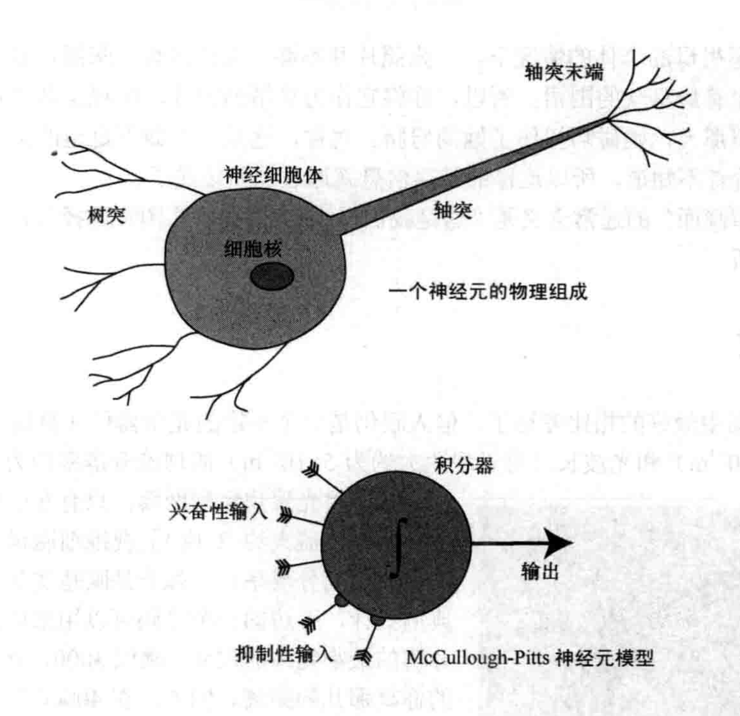


图 2.3　警方草图与嫌疑人被逮捕后所拍摄面部照片的比较

一个阈值逻辑单元基于许多输入的加权和的实现，可以进行信号识别的过程。该过程并不复制一个实际神经元的确切功能，但基于 McCullough-Pitts“感知”，这成功地描述了整个过程（见图 2.4）。这一概念将在第 12 章中作为实现相似分类的软件模型再次出现。

识别经常用“祖母细胞”来描述。这是一个理论结构，而不是在大脑中某处的单个物理细胞，但它提供了一个有用的框架去描述识别过程中的一些显著的特征。祖母细胞的思想是，它耐心地检查每个图像寻找“祖母”的外观，然后向意识发出信号表明她存在。对到达眼睛的原始图像的处理在几个地方发生，包括视网膜和视觉皮层，并以一种并行的方式进行。在这一过程中，可以粗略描述为颜色、大小、位置和形状几个特征并被提取出来。其中一些可与祖母的存储模型（如身材矮小，白发，一个微笑，甚至可能是一件熟悉的礼服）相匹配。显然，有些特征比其他特征更重要，所以输入就必须有权重。如果存在足够的正确匹配，在没有强烈加权的反面特征（如红胡子）时，就会发出“祖母”信号。

这个“阈值逻辑单元”的简单模型演变成了现代的神经网络，其中那些独立决策的几个层使得单元连接在一起。它们的输入将来自不同传感器和来自其他单元的输出的数据组合起来，并最终输出一个决策，该决策基本上是识别输入与一些识别出的环境或目标（足够好地）匹配。神经网络决策的特征，无论是在人类思想中的图像上，还是在计算机电路中其他类型的数据上，都具有非常高的速度（归于极端并行的方式，所有的小逻辑决策在同一时刻发生）、很好的学习能力（通过调整加到不同输入的权重）以及犯错误的倾向。



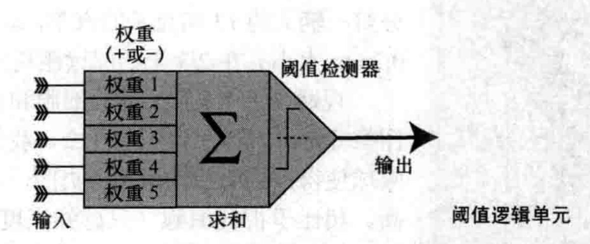


图 2.4 物理神经元、神经元的 McCullough-Pitts 简化模型及其阈值逻辑单元实现的比较。如果多个输入的加权和超过了某个阈值，则打开输出（它连接到另一个逻辑单元的输入）。学习是调整权重的过程，权重可正可负

每个人都有这样的经历，即认为他们认出了某人（“祖母”），然而仔细观察后意识到，所认出的人完全不是那个人。有足够的正确线索，而没有反面的线索，来引发识别过程。也许在不同的情况下，或以不同的观点，我们将不会犯那样的错误。将正确输入的加权和的阈值设置得过高，将会减少错误判断，但效率较低，需要太多的时间去收集更多的数据。做出错误识别的惩罚有点尴尬。快速、高效的程序的好处在于，能够进行基于不完整数据的识别。

在该逻辑的一些实现中，可以对某个辨别安排一个概率或置信度，但该值的作用在很大程度上依赖于底层模型的质量。这可以表示为神经网络的权，或一个模糊逻辑系统的规则，或是某些其他形式。在识别中，模型中的因素列表并不是那么明确。模型写下所有的特征来帮助辨别祖母（尤其是反面排除）是非常困难的。在大多数科学实验中，一项工作是罗列重要因素，但总有潜在假设的背景水平，它可能也可能不会被那些查看结果的人分享。

在涉及成像的科学论文中呈现出一幅图片是很普遍的，通常以“典型的外观”或“代表性的观点”为标题。科技期刊的编辑了解这些照片的目的是为了显示模型列表中的几个因素，这几个因素是作者认为非常重要（并希望在文本中描述）的。但没有图片会是真正“典型的”。对于一件事，最自然发生的结构有一些变异，并且发现一个个体中的所有特征的均值的机会是很小的，而且在任何情况下不会展现出变化的范围。

但是，即使在像祖母的个体的情况下，一张照片并不够。也许你有一张照片显示她的脸、白发、无框眼镜，她甚至穿着最喜欢的围裙。所以，你将它作为典型的图片，但观看者注意的却是她的一条臂悬带，因为你拍照那天，她碰巧扭伤了她的肩膀。对你，这是一个微不足道的细节，而不是图片中的重要事情。但观看者不知道，所以这样的错误信息就通过画面传送了。

编辑知道“典型画面”的通常含义是“这是我们最漂亮的图片。”图片选择经常包括一种偏离图像的许多应用审美判断。

2.3 技术参数

尽管与动物王国中最好的相比差远了，但人眼仍是一个不错的光学器件（见图 2.5）。基于晶状体孔径最大尺寸（5×10^{-3}m）和光波长（对于绿色大约为 5×10^{-7}m）的理论分辨率约为 10^{-4} 弧度或 1/3 弧分。晶体将光聚焦到视网膜，只有在中央凹［视网膜的一小部分（覆盖大约 2 度）］视锥细胞最为密集，得到感知图像的最高分辨率。一弧分是眼睛实际分辨率性能的一个典型估计，手边的一个号码可以用来估计距离。估计你能分辨的最小物体的尺寸，乘以 3000，就是以同一单位表示的你离那儿的距离。例如，在 40000 英尺高的飞机上可以分辨一辆大约 13 英尺长的汽车，本书中的细线和空白（约 0.2mm 大小）在 2/3 米的阅读距离处可以分辨等。

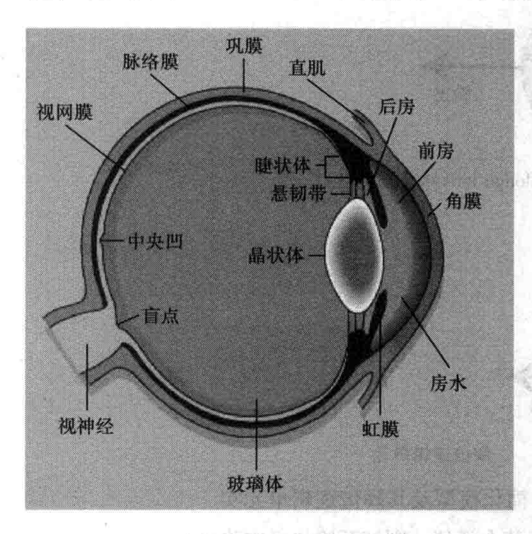


图 2.5 眼睛的简化图，显示了晶状体、视网膜、中央凹、视神经等

视网膜上 1.6 亿个视杆细胞和视锥细胞不估计实际的图像分辨率。为“看到”什么，我们旋转头部或眼眶中的眼球使得该点的图像落到中央凹，那里的视锥细胞密度最高。相比于视杆细胞（只感知亮度），视觉的周边有相对较少的视锥细胞（对应于颜色），其最重要的作用是检测运动和判断场景光照来校正色彩平衡与遮蔽。要生产一台捕获整个视觉场景（也就是要求一个有非常广角的镜头）且其分辨率与人类中央凹视觉相匹配的数码相机，来查看所存储图像的任何位置及查看原始场景中已可视的最精细的细节，需要数十亿个传感器。

人类视觉在没有任何意识的情况下，可通过快速移动眼睛来查看一个场景中的许多不同位置，并将这些零散的内容组合成单幅可感知的图像，以实现一些非常奇妙的事情。即使当我们“看”某个东西时，眼睛也不是静止不动的。通过不断地移动眼睛，人们在头脑中会构建一幅场景。在视网膜上有一个盲点，该处没有视神经连接视杆细胞或视锥传感器。我们并不会注意到盲点的存在，因为大脑会根据周围的场景或记忆片断来进行内插。物体在盲点出现或消失的测试表明，我们实际上从那里并没有获得任何信息，所获得的信息只是大脑为我们编造的。

眼睛在非常宽的光照级别范围内都可以捕捉到画面，横跨 9～10 阶幅度，从星光闪烁夜晚的数十个光子到滑雪坡上明亮的晴天。有些适应源于虹膜孔径的改变，但绝大部分适应取决于视网膜中的处理。光照级别改变的适应需要花一些时间，最多可达几分钟，时间的长短取决于变化量。在幅度最暗的几阶，只用到了视杆细胞，而不能分辨颜色。在黑暗中，尽管中央凹处的视锥细胞很丰富，但由于几乎没有视杆细胞，因此较好的策略是看“相邻的”目标（“眼角余光法”）。这种方式会使用具有更多视杆细胞的区域来捕获昏暗的图像，虽然分辨率要差一些。

视杆细胞对可见光谱的红色一端不是非常敏感，这就是为什么天文学家、潜水员和其他希望能够关闭红光并在视杆细胞视觉中具有全敏感性的这些人采用红光照明的原因。视锥细胞有三类，每类对应于

稍微不同的波长范围（见图 2.6）。它们通常被称为长波长、中波长和短波长受体，或更加简洁但不那么准确地称为红、绿、蓝敏感。通过比较每一类视锥细胞的反应，眼睛对颜色进行描述。例如，黄色是红色和绿色的组合，品红相对缺乏绿色等。灵长类视锥三色组的进化与其他动物的进化树不同（Jacobs & Nathans，2009），且与其他动物（如鸟类和章鱼）眼中的彩色传感细胞相比，对应于稍微不同的波长。

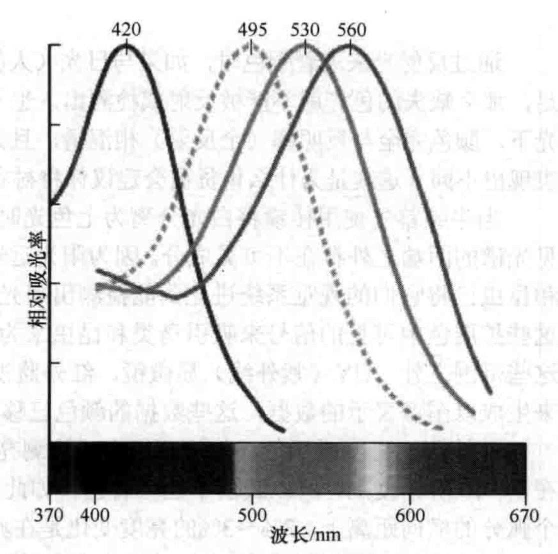


图 2.6　视杆（灰色）和 3 种视锥（红色、绿色和蓝色）的敏感性与波长的关系（Dartnall et al.，1983）。人类视觉检测的范围大致是 400nm（蓝色）～700nm（红色）

因为红、绿、蓝敏感视锥有不同的密度，眼睛对绿光反应最强而对蓝光反应最差［如第 1 章的图 1.12(b)所示］。但是，这种敏感性是有代价的：在相同的绿色波长范围内，我们对颜色的分辨能力是最弱的。由于眼睛的敏感性，显微镜中的一种常用技术是使用滤镜来仅选择绿光，但如果颜色变化的检测非常重要，这就不是一个好的策略。

像眼睛所做的大多事情一样，对颜色的感知通过比较而非绝对的方式确定，这会产生许多对颜色不正确的解释（Werner et al.，2007）。只有通过与已知彩色基准相比较，我们才能真正地估计颜色。常用的彩色基准是一个定义了所有颜色的白色物体。如果所看的景象中包含有中性灰色的物体，那么光照颜色的变化可以对其进行补偿（如第 1 章中所述）。它不像看起来那么简单，因为许多物体不能同等地反射出所有颜色的光，而会随着角度或光照出现颜色变化（这种“同色异谱”现象也是喷墨打印机经常出现的一个问题）。

因为只存在三种对颜色敏感的视锥受体，且每一种受体都有着宽的和重叠的波长响应，所以许多不同的波长组合会引起相同的视觉响应，且这种视觉响应会因个体的不同而不同。将在第 3 章中详细讨论的色彩匹配是一个需要仔细校准的专业话题，因为色彩匹配会在打印图像与显示图像时用到。

一个常见的问题是：“RGB 的哪些部分对应于光的一个特定波长？”由于某些原因，这个问题并没有答案。感知的颜色依赖于人类的感知，且在一定程度上因人而异，还依赖于观看的环境。三类视锥敏感性的较宽（重叠）波长范围表明，同样的颜色感知可由波长和亮度的许多不同组合产生。由于荧光粉不同（以及它们的老化），以及 CRT 与 LCD（液晶显示器）之间的差异（取决于背景照明的颜色），使用 RGB 分量在不同的电脑显示器上生成颜色是非常困难的。然而，表示颜色随波长变化的一种有用方式是，使用如图 2.7 所示的模型（Bruton，2005）。对于印刷而言，这一问题（将在第 3 章中讨论）会更为严重，因为该问题也取决于图像在什么样的纸张上以及查看图像时的光照条件。

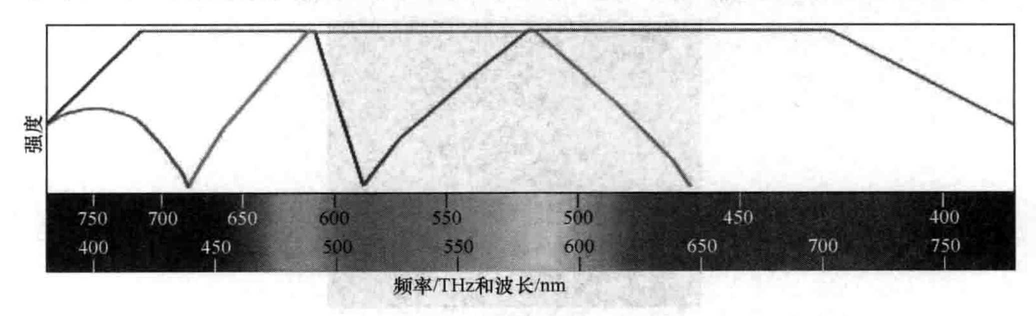


图 2.7　波长与计算机显示所用 RGB 亮度间关系的一种近似视觉表示

通过反射光来观看颜色时，如果与日光（人类视觉在阳光下逐步形成）相比光照在色谱某些部分不足，那么缺失的色彩就不能被反射或检测出，也不可能准确判断出目标及其真实的颜色。在单色的黄钠光下，颜色完全与反照率（全反射）相混淆，且无法由亮度来分辨。即使在典型的室内光线下，颜色也表现出不同（这就是为什么售货员会建议你将衬衣和领带拿到窗口去看它们在阳光下的颜色的原因）。

当牛顿首先使用棱镜将白光分离为七色光时，还通过测量由吸收该光引起的温度上升，证明在可见光谱的两端之外存在不可见成分。因为阳光远远超出了400～700nm的人类视觉范围，所以一些动物和昆虫已将它们的视觉系统进化到能检测出阳光，就不会让人觉得惊奇了。植物尤其会发出一些仅在这些扩展色中可见的信号来吸引鸟类和昆虫来为其授粉。通过某些仪器，甚至可以将人类视力延伸至这些范围之外。UV（紫外线）显微镜、红外监视、射电天文学等都利用了可见电磁光谱之外的部分，来生成以图像表示的数据，这些数据的颜色已移到我们能检测的谱的较窄部分。

能检测出亮度或颜色，与能测量它或检测亮度或颜色中的较小变化是两回事。因为人类视觉工作在9～10阶幅度上，所以我们不能查看包含如此大范围的单幅图像，也不能检测出$1/10^9$的变化。在几个弧分的横向距离上，2%～3%的亮度变化是在典型观看条件下可检测的极限。必须注意，所要求的变化是一个百分比，所以在亮区比在暗区更需要亮度上的更大的绝对变化。任何具有摄影暗室经历的人都能意识到，不同的细节在负片上比在正片上更容易看到。

总体来说，眼睛只能检测出一幅图像中的20～40种灰度，许多情况下，更少的灰度种数能产生视觉上更令人满意的结果。在一幅主要由不同亮度的大区域组成的图像中，很难挑出每个区域中具有较小局部对比度的精细细节。提高局部细节可视能力的一种常用方法是，减少亮度的全局（长范围）变化而增大局部对比度的计算机增强。这通常通过比较一个像素与其局部邻域来完成。若像素比其邻域更亮，就使它仍然更亮，反之亦然。这是第5章中所展示的某些增强技术的基础。

人类视觉通过对局部区域的自适应及检测局部对比度来处理包括一个很大亮度范围的场景。这些景物无法通过胶片或数码相机（它们都具有有限的动态范围，最大在3～4阶幅度上）来整体记录。通过用不同曝光来记录这样的场景并通过合适的软件将它们组合起来，就有可能通过牺牲大范围的对比度来保留局部对比度，如图2.8所示（曝光合并过程将在第5章中介绍）。第3章中显示了处理高动态范围图像的相关方法，以便图像可以在降低总体对比度的同时保留局部细节和对比度的情况下显示或打印。

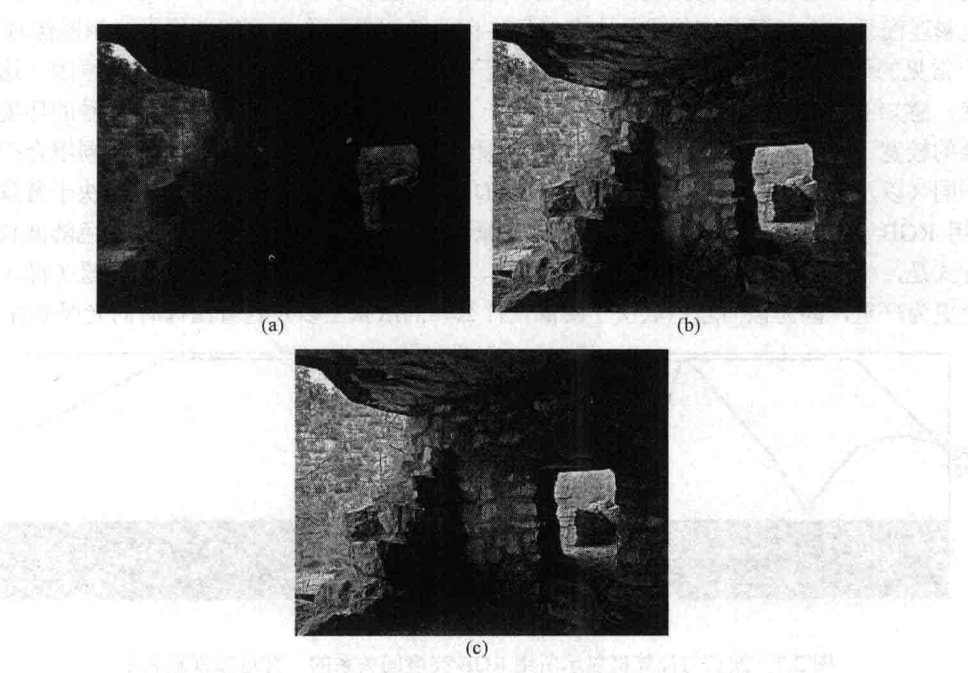


图2.8　不同曝光时间的图(a)和(b)合并后形成了图(c)，保留了局部对比度

亮度（或颜色）的局部变化或突然变化是图像中最容易注意到的细节。纹理中的变化通常表示重要的结构变化，但这些变化更加精细。如图 2.9 所示，在许多情况下，分类为纹理的变化实际上表示平均亮度级的改变。当纹理中只有一个变化存在时，视觉检测很困难。如果边界线不是直线（尤其是垂直和水平方向），将更难看到。第 5 章包括了基于纹理差异来分辨区域的方法。

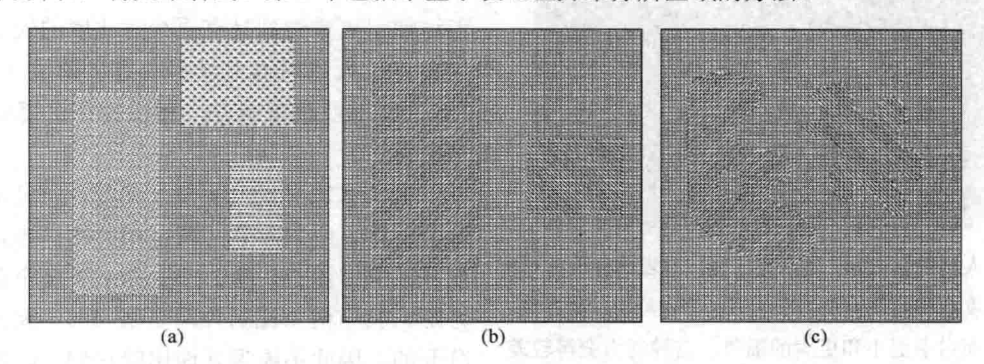


图 2.9　纹理差异示例：(a)平均亮度不同的区域；(b)无亮度差异但有简单的线性边界；(c)不规则边界

红、绿、蓝视锥各自的 30 级亮度表明可以分辨 30^3 = 27000 种颜色，但事实并非如此。在谱的两端对颜色变化的敏感度比在中间的更高（换句话说，绿色很难从其他颜色之中区分出来）。只有大约 1000 种不同的颜色可以区分。由于计算机显示器为 R、G 和 B 提供 256 级亮度，或 256^3 = 1600 万种颜色，可以预料到它们可以生成我们可以看到的任何颜色。然而，情况并非如此。和我们所能看到的相比，计算机显示器和印刷图像都受到色域（能产生的颜色的全部范围）的限制。这是"典型的图像"并不能成为对象或对象类的实际表示的另一个原因。

2.4　敏度

许多动物，尤其是鸟类，能够产生比人类更高空间分辨率的视觉。人的视觉仅在视场中心的一个很小区域（中央凹）达到其最高空间分辨率，在此处，感光视锥的密度是最高的。在 50cm 的观看距离处，宽约 1mm 的细节表示稍微超过 1/10 度的一个角度。敏度（空间分辨率）的单位通常为周期/度。人眼可见的上限（最精细细节）大约为 50 周期/度，这对应的一个光栅，其亮度在同样的 1mm 内从最小到最大变化大约 5 倍。在这个精细的间隔中，需要 100%的对比度，换句话说，即黑线和白色空间。这是共同规范出现的位置，即在没有光学器件帮助的情况下，可分辨的最精细的线大约为 100μm（在这种情况下，"可分辨"表示可看到 100μm 宽的线，且这些线被 100μm 对比空间分隔的线分开）。

当特征更大时，检测这些特征时在亮和暗位置所需要的对比度更小。在正常的阅读距离，大约 1mm 宽的亮度变化表示一个约 9 周期/度的空间频率，且在理想的观看条件下，可以以百分之几的对比度进行分辨，但要假定图像中没有任何噪声且是一幅非常明亮的图像（在暗图像或叠加随机变化的图像中，敏度会显著下降，比起检测亮度变化而言，检测颜色差异的能力更差）。

在约 50cm 的正常观看距离处，图像上的 1mm 约是检测所存在细节的最佳尺寸。在典型的计算机显示器上，这大约相当于 4～6 个像素。随着空间频率的下降（特征变大），所需的对比度增加，所以当亮度从最小变到最大的距离大约是 1cm 时，所需对比度大约增大 10 倍。空间分辨率（敏度）随对比度的变化（Barten，1999）被称为调制传递函数（见图 2.10）。

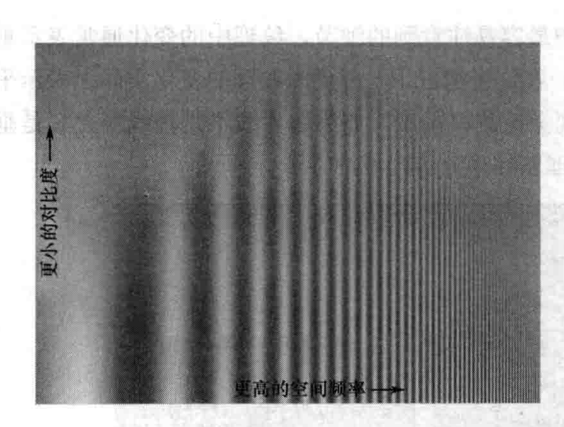


图 2.10　人类视觉调制传递函数示例，它表明分辨低对比度细节的最大能力出现在一个中间空间频率处，而对于更小和更大的细节，这种能力变得较差

放大图像并不能提高分辨小细节的能力，相反会经常退化图像。显微镜使用者常犯的错误是，以非常高的放大倍数工作，并期望看到最精细的细节。对于小尺寸的细节来说，这是需要的，但这使得更难看到具有更少对比度的较大特征。同理，放大计算机上显示的数字图像（“像素检视”）并不能改善图像，相反往往会降低看到细节的能力。

因为人眼不能“测量”亮度，而只是进行简单的比较，所以除非区域是相邻的，否则分辨亮度差异非常困难。图 2.11(a)显示了 4 个灰色的正方形，其中两个比另外两个暗 5%。因为它们是分开的，因此不能很好地比较它们。即使是区域相邻，如图 2.11(b)所示，如果从一个区域到另一个区域的变化是渐进的，也很难察觉。仅当灰度突变时，眼睛才能判断哪些区域是不同的，如图 2.11(c)所示。

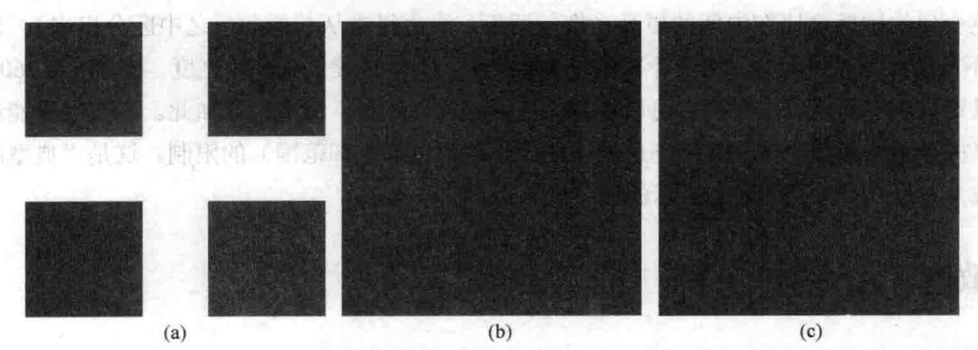


图 2.11　具有 5%亮度差异的区域的比较：(a)分开的区域；(b)相邻但有渐进变化的区域；(c)相邻但具有突变边界的区域

图 2.12 显示了几组方块。在每一组中，有两个方块是相同的，另一个不同。我们使用相同的方式来测试“刚好能察觉到的变化”（Matacara-Hernandez，2002）。第 1 组中，块 C 比其他两个块浅 5%。第 2 组中，块 B 的色调向绿色偏移了 5°。第 3 组中，块 A 的色调向蓝色偏移了 5°。第 4 组，块 C 的色调向蓝色偏移了 5°。需要注意的是，这些图片是用于在计算机显示器上观看的，图片间的差异会因本书的印刷技术增大或减小。

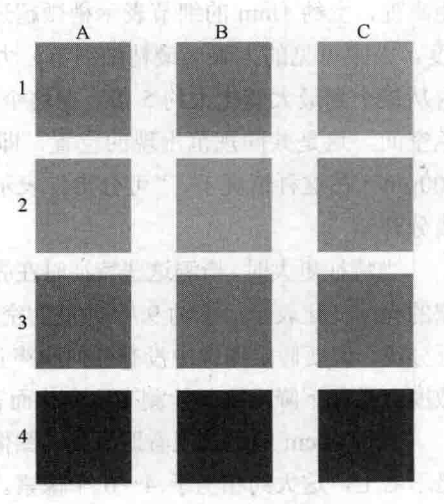


图 2.12　文中所述用来测试亮度或颜色“刚刚能察觉到的变化”的色块

即使当特征有应该可见的尺寸和对比度时，背景亮度（或颜色）的变化也会阻止视觉系统检测出它们。在图 2.13 的例子中，文字有一个约 2%的局部对比度，但它叠加在从白到黑变化的过渡上。图像处理运算的应用揭示了这一消息。“反锐化模板”程序（详见第 5 章）从原始图像减去了一个模糊（平滑）的副本，抑制了大尺度的变化，从而显示出了局部细节。

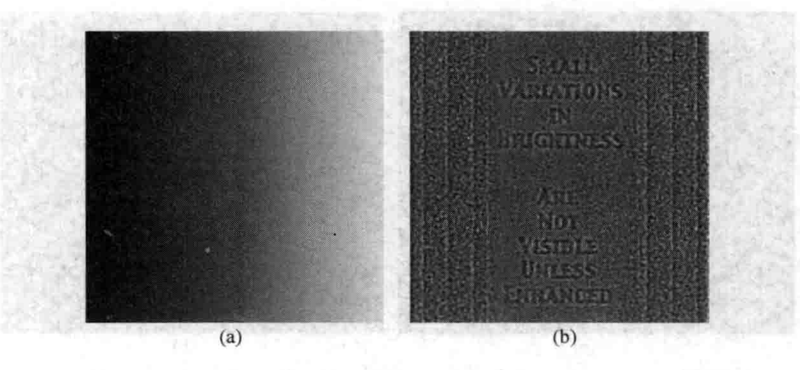


图 2.13　叠加到变化背景上的亮度差异在视觉上是察觉不到的：(a)原始图像；(b)使用“反锐化掩模”滤波器抑制渐进式变化并揭示细节后的图像

分辨率和能见度这两个概念很容易让人混淆。天空中的星星基本上是一个点，没有角大小，甚至在望远镜中绝大多数星星也不会看上去像一个盘子。星星因其与黑暗的天空相比有着较高的对比度，所以是可见的。暗星并不一定更远或更小，但因为没有足够的对比度，所以肉眼看不见。望远镜通过更大的光圈收集更多的光线，可使它们可见。考虑分辨率的一种更好方式是，区分或分开靠得很近的两颗星星的能力。长期使用的一种经典测试是北斗七星斗柄上的开阳星（Mizar）。在凡・高的作品“满天星斗下的罗纳河”中，该星群中的每一颗星都被显示为单个实体（见图 2.14）。

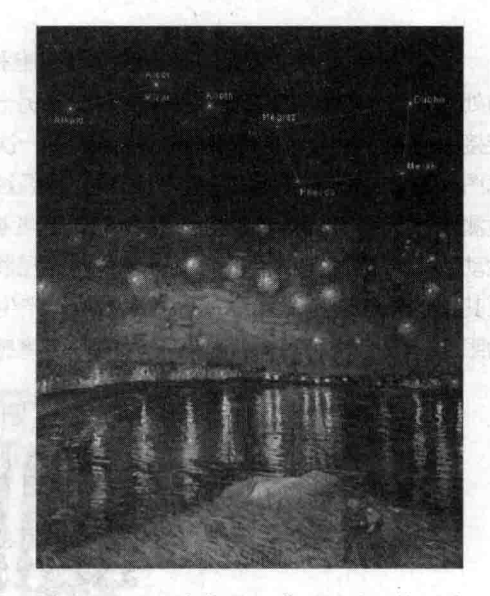


图 2.14　凡・高作品“满天星斗下的罗纳河”上的北斗七星与星图的对比

但星图表明，北斗七星斗柄中的第二颗星实际上是双星。辅星（Alcor）和开阳星是光学双星——两颗星看上去靠在一起，但它们与我们的距离并不相同，且它们之间没有引力。它们相隔约 11.8 弧分，美洲的印第安人和中东的沙漠居民（在没有紊流或水蒸气，且没有月亮或其他光污染的黑暗天空下），能够将这两颗星分辨开。

人们对开阳星的了解实际上更多。随着伽利略望远镜的出现，天文观察者惊奇地发现，开阳星本身就是一个双星。通常认为，博洛尼亚耶稣会的天文学家和地理学家 Giovanni Battista Riccioli（1598－1671）大约在 1650 年就分开了开阳星——所发现的第一个双星。双星开阳 A 和开阳 B 是相隔 14.42 弧秒的引力双星，任何良好的现代望远镜都可以分开它们［见图 2.15(a)］。但它们被证实更加复杂——每颗星本身也是一个双星，它们靠得很近，只有通过光谱学才能检测出它们。

最近的发现表明辅星也是双星。辅星有一个红矮共生辅星 B，其质量只有辅星 A 的 1/8 左右，且只有用日冕仪并遮住来自更亮星星的光时才能看得［见图 2.15(b)］。天空中有许多双星可以用后院望远镜分辨出来，其中许多是业余天文学爱好者所熟悉的观察目标，他们的目的是找出一些星对的颜色差异，或分辨出它们。这不仅仅取决于角分离。如果一颗星星明显要比其他星星明亮，那么在亮星附近是很难看到暗星的。

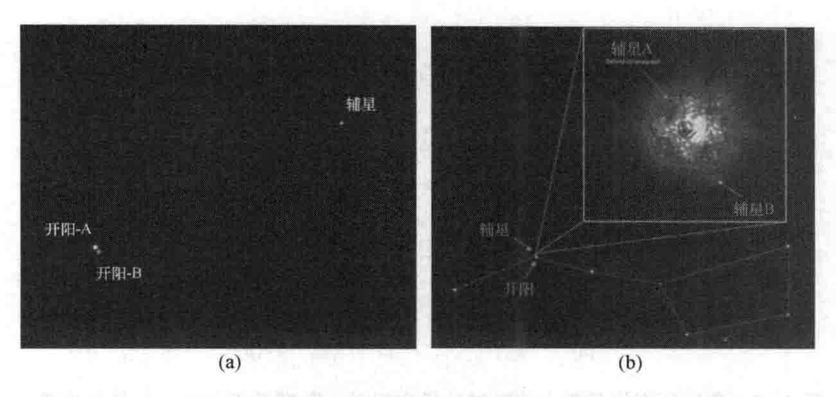


图 2.15 双星：(a)显示辅星与开阳 A 和开阳 B 的相对分离的照片；(b)遮住来自开阳 A 的光可检测出伴随开阳 B 的更小红矮星

2.5 眼睛告诉大脑的内容

人类视觉远不止是视网膜上的视杆细胞和视锥细胞。在一幅“图像”存在于意识之前，发生了大量的处理，一些处理立即发生在视网膜上，另一些处理则发生在大脑后方的视觉皮层上。视网膜中的神经连接大约在 100 年前由 Ramón y Cajal 第一次看到，并一直被研究至今。图 2.16 显示了人类视网膜的简化图。感光视杆和视锥在后方，光线必须经过几层处理细胞才能到达。进入人眼的光子中只有约 10%被检测到；其余的或者被反射（约 3%），或者被无意义的结构吸收，或者落入活性分子之间。在许多夜行性动物中，在视杆和视锥后面有一个色素层将光线反射回来，使得将要捕获和检测的光子达到两倍（而且其中的一些通过晶状体折射回来，从而产生我们所看到的夜间注视我们的反光眼）。顺便说一句，章鱼的眼睛并没有这种后向排列；此时的进化将感光细胞置于上方，在那里它们可以最有效地捕捉光线。

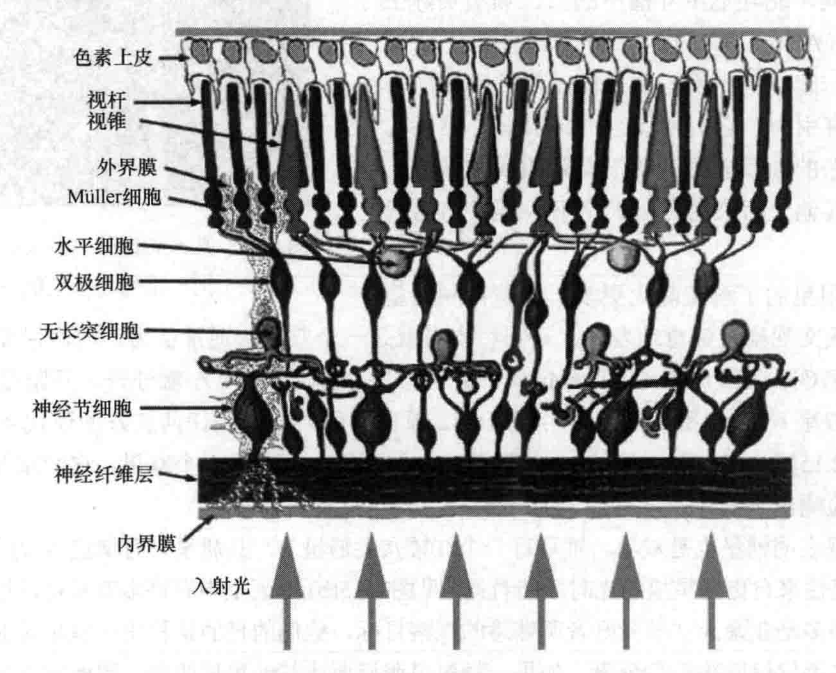


图 2.16 视网膜中的各个主要层。光线穿过几层处理神经元到达对光敏感的视杆细胞和视锥细胞。水平细胞、双极细胞和无长突细胞将来自不同大小区域的信号进行组合，比较它们以便找到感兴趣的细节，并将这些信息传递到视觉皮层的更高层

处理细胞的第一层，即水平细胞，连接相邻的不同大小的感光细胞。双极细胞传输来自这些感光细胞及来自下一层的感光细胞的输出，在下一层中，无长突细胞将信号组合并比较。最后，神经节细胞收集这些传输的输出到视觉皮层。在视神经中的传感器数量大约是神经连接中的 100 倍，这意味着需要大量的处理来提取有意义的信息。这一物理组织直接对应于抑制的逻辑处理，讨论如下。

在许多方面，眼睛的视网膜是大脑的一部分。理解图像数据的前期处理及从眼睛传送到视觉皮层的信息的提取的研究被授予诺贝尔奖[③]（1981 年，David H. Hubel 和 Torsten N. Wiesel 因其关于视觉系统的信息处理的新发现而获奖）。

若没有详细说明视网膜的解剖细节，或由 Hubel、Wiesel 发现的视觉皮层，或其他信息（尤其是由 Jerome Lettvin 发表于 1959 年的一篇开创性论文"蛙眼告诉了蛙脑什么"），仍然可以总结出其逻辑和实践意义。在视网膜内，来自独立的感光器的输出被各神经元层组合并比较。将来自一个传感器或区域的输出与来自周围传感器的输出进行比较，测试中心的激励与来自周围的抑制，是使视网膜忽略均匀的或仅在亮度上渐变的区域，并有效地检测亮度发生变化的位置的一个基本步骤。对不同大小区域的测试就定位了不同尺寸的点、线和特征。

以同样的方式可以比较不同时间的输出，进而检测变化，过快或过慢的变化都看不到。尽管昆虫和一些鸟类可以注意到荧光灯 60Hz 的闪烁，但人类不行。对超过几十秒的任何变化进行调整，可防止这些变化被注意到。

在蛙眼中，视网膜会处理图像，以便找到一些非常具体的激励。这些激励包括小而暗的移动目标（食物）和最大、最黑暗区域的位置（在池塘中安全）。如苍蝇和人类一样，蛙眼会"隐约可见地"去检测那些迅速增大且在视场中并不移动的东西。这代表了朝眼睛而来的东西，并使得苍蝇避开苍蝇拍或青蛙跳入水中。而人类会在意识到一个对象甚至看到对象之前，有意识地眨眼来保护眼睑。

接下来在大脑皮层进一步组合这些原始检测回路的输出，进而找到线和边缘。大脑皮层中的一些特定区域对不同方向的线及其运动较为敏感。这些区域的"连接"不是与生俱来的，但一定会在出生后演化；在缺乏特定方向的线的环境中饲养的小猫，以后没有能力看到这样的线。检测亮度变化的位置（特征边缘）会产生一种场景的心理草图，其中线、边缘、角及其他简单的结构占主要部分。这些反过来在层次结构中联系在一起，进而在我们的大脑中产生场景的理解。"图像理解"软件试图对场景图像做同样的事情，以便构建场景几何模型，进行目标识别及基于由边缘形成的草图的定位。

对亮度或颜色随位置或时间变化的提取，很大程度上解释了场景中什么被感知到，什么被遗漏。随位置逐渐发生的任何变化（如墙壁上光的阴影）常常被忽略。我们必须尽力去注意这样的阴影，即使是在没有量化工具来估计其幅度的情形下。但当它们突然发生时，即便是在亮度上几个百分比的很小变化也都是可见的，从而产生确定的边缘。边缘形成一条直线（尤其是当它接近垂直）时，它会被注意到。同样，随时间保持静止的场景的任何部分都很容易被忽略，但当某些东西移动时，它就会吸引我们的注意。

因为从场景中提取信息的这些技术是高度并行的，所以非常高效。对于眼睛中的 1 亿 6 千万个感光器中的每一个，都有多达 50000 个神经元参与处理和比较。Hubel 和 Wiesel 的贡献之一是，展示了在出生后不久网络连接是如何形成的，以及对这种形成在关键时期给眼睛提供不同图像的关联性。

大脑中具体电路的映射，是通过在大脑皮层的不同位置放置电极并观察神经元以各种图像和呈现给眼睛的刺激作为输出的过程所完成的。在较高尺度的处理中，功能磁共振成像和 PET 扫描可以识别大脑对不同活动和刺激的反应变得活跃的区域。但是在脑海中的图像处理还有另外一个重要的知识来源：识别图像解译中造成的错误。一个重要但有限的资源是正在研究的具有特定损伤的人的反应，这类损伤可以是天生的，或者是由疾病或事故造成的。第二种方法研究由视觉上的错觉造成的图像解译

③ 生理学或医学奖。——译者注

错误。既然每个人都倾向于犯同样的错误，那么这些错误必定是处理如何完成的一个直接指示。一些更具有揭示性的情况会在后面的章节中给出。

2.6 空间比较

中心环绕或激励抑制逻辑背后的基本思想是，将来自中心区域的信号（可能是一个单一的检测器，或多个检测器的平均）与来自周围环形区域的输出进行比较。这是一个在视网膜上的基本分析单元，通过把许多这样的原始单元的输出组合在一起，可以检测亮线或暗线、角、边缘及其他结构。业已证明，一定大小的暗色物体（对应于近到可以被捕获的昆虫）会在蛙眼中产生强烈的识别信号。在猫的视觉皮层中，存在仅响应一定长度和角度并以某个特定方向移动的暗线的区域。这些区域与来自另一只眼睛对图像数据执行同样分析的区域相交汇，这正是立体图像融合的基础。

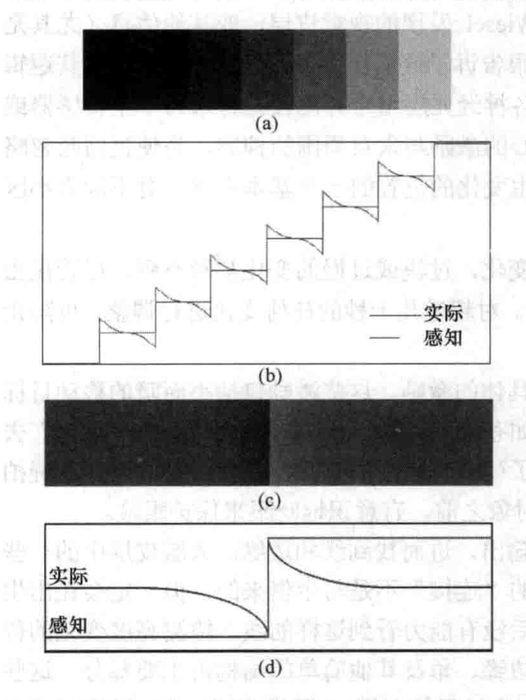


图 2.17　基于抑制的两种常见错觉：图(a)、(b)马赫带表明视觉系统在各阶处增大了亮度的感知变化；图(c)、(d)Craik-Cornsweet-O'Brien 阶表明，为视觉系统提供一个阶会影响到对更远的值的判断

这一基本的中心环绕行为，解释了几种很常见的错觉（见图 2.17）。一组均匀灰色阶（马赫带）未被感觉成均匀的。因为注意到的是阶，所以每个阶靠近较浅区域的一边被感知得更暗，反之亦然。就马赫带而言，对阶的视觉反应有助于我们确定哪个区域更暗，尽管这会使得我们很难判断差异的大小。但同样的效果可用于在没有任何差异的情况下来使得物体被察觉。在 Craik-Cornsweet-O'Brien 错觉中，两个邻近的区域有相同的亮度，但亮度值在边界的一侧增大而在另一侧减小。眼睛用与马赫带相同的方式来解译它，而且判断出一个区域较浅而另一个区域较深。

类似的激励抑制比较用于颜色值。色块间的边界被检测并加强，而色块间的绝对差异被最小化。放置在亮度或其他一些颜色的梯度上的相同色块看上去是不同的。

几个计算机图像处理算法使用了中心环绕，包括反锐化模板和第 5 章中给出的顶帽滤波器。此处给出的一些寻找边缘算法的实现，也模拟了视觉系统中进行的处理。

不管是以亮度、色调和饱和度，还是以艺术家的色彩、色度和色调，或是以各种数学空间进行描述，色彩都需要三个参数。第 1 章中给出的所有色彩空间坐标系使用三个值来定义颜色。亮度是光的亮或暗的度量，而不考虑任何色彩信息。对艺术家来说，这就是色度，通过往颜料中加入黑色来实现。色调区分各种颜色，围绕我们可能在幼儿园见过的色轮，从红变到橙、黄、绿、蓝和品红。饱和度是呈现色彩的数量，如粉红色和红色，或天蓝色和品蓝色之间的差异。全饱和的色调对应艺术家的纯颜料，可加入中性颜料来降低饱和度，进而增色（色调描述向纯色中加入白色和黑色颜料）。

即使没有色彩学（或艺术）的正规培训，这也是人们向自己描述颜色的方式，虽然未必使用相同的术语。组合和比较来自对红、绿、蓝敏感的视锥的信号，可确定这些参数。简单地说，我们可以认为所有三个参数的总和，或对红、绿和蓝检测器不同敏感程度的加权和，就是亮度；而一个参数与另一个参数的比被解释为色调，同时最大参数与最小参数的比对应于饱和度。记住，色调不直接对应于

从红到蓝的波长范围。我们将具有弱绿色、强红色和蓝色的颜色解释为品红色或紫色，这不是一种波长意义上的颜色，但肯定是根据感知而来的。

依赖于比较而不是绝对值来检测亮度或色彩的变化，简化了许多图像解译任务。例如，一栋建筑物的不同墙壁都涂成相同的颜色，因为遮蔽、与太阳所成的角度等原因，可能有非常不同的亮度值，但因为仅用于计算局部变化，我们所观察到的是良好角落和细节的颜色均匀的建筑物。图 2.18 所示建筑物阴影侧的“浅灰色”墙壁实际上要暗于阳光照射侧的“黑”百叶窗，但我们的感知认为该建筑物的所有边和墙壁是浅灰色的，门窗镶边是白色的，而百叶窗是黑色的。

如图 2.19 所示，局部比较的视觉依赖性意味着它不可能确定场景中某一点或区域的绝对亮度。同一局部比较原理适用于许多其他情况。许多艺术家一直使用类似于 Craik-Cornsweet-O’Brien 的错觉，来使画布上的画的亮度绝对值局限于非常有限的范围内，进而创建很大范围的光影效果（见图 2.20）。通过增加颜色，例如使用诸如绿色或品红色这样没有物理基础的颜色来对皮肤上的阴影着色，可增强这一效果，但同时会增强对亮度差异的感知（见图 2.21）。



图 2.18　正如文中讨论的那样，建筑物阴影侧和阳光照射侧的亮暗区域的比较在局部进行

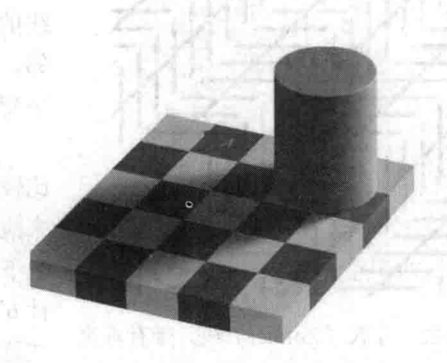


图 2.19　亮度的视觉解释受期望值及周边环境的影响。我们“了解”阴影，并能自动补偿其对亮度的影响，同样，我们“了解”棋盘，故标记为 A 和 B 的方块可以分别理解为黑色和白色，但它们实际上具有相同的灰度

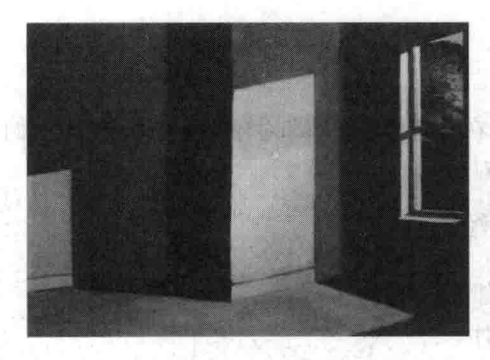


图 2.20　爱德华·霍普（Edward Hopper）④画了许多光和阴影的效果图。其作品“空房间中的阳光”显示了沿均匀表面和边缘台阶的阴影



图 2.21　眼影（和其他化妆）形成了许多色调

宝丽来知名人士 Edwin Land 广泛研究了颜色视觉，并给出了与前面描述的三刺激模型（这通常归功于 Helmholtz）稍微不同的解释。高度依赖于用于处理信号的激励-抑制模型，Land 的对抗色“Retinex”

④ Edward Hopper（1882 年 7 月 22 日—1967 年 5 月 15 日），美国绘画大师，以描绘寂寥的美国当代生活风景闻名。——译者注

理论预测或解释了一些有趣的视觉现象。注意，在 Land 的理论中，来自一幅图像中某个区域的合成光本身并不能指定该区域的感知颜色，相反，该区域的颜色由三个数来确定，在单个波段（长、中和短波长通常粗略描述为红、绿和蓝）上计算出的数，给出了场景中该区域与其他区域之间的关系。

Land 理论的结果之一是，光照谱成分变得非常不重要，一个特定区域的颜色可以变得与照射其的光相对无关。人眼有相对较少的视锥去感知视觉边缘的颜色，但该信息通过校正入射光的颜色，在判断中心的颜色时显然很重要。Land 证明，若场景由红色、绿色和蓝色滤镜拍摄，且随后使用投影仪加亮了刚好通过两个负片的彩色光或白光，则场景可被感知为全彩色。另外一个有趣的现象是，有着黑白条纹的转盘可被感知为有颜色的，具体取决于条纹的间隔。许多关于如何在人类视觉系统中精确处理色彩信息的问题至今仍然没有答案。

上面关于中心环绕比较的讨论主要集中于视网膜中的处理，它将每个点与其周围进行比较。但由于提取的信息提升了处理链，通过视觉皮层，存在许多其他局部比较的证据。如上所述，视觉皮层中存在响应于特定角度线的一些区域。它们的位置与响应稍微不同角度线的区域相邻。比较一个区域的输出与其邻域的输出，可检测角度上的较小变化。

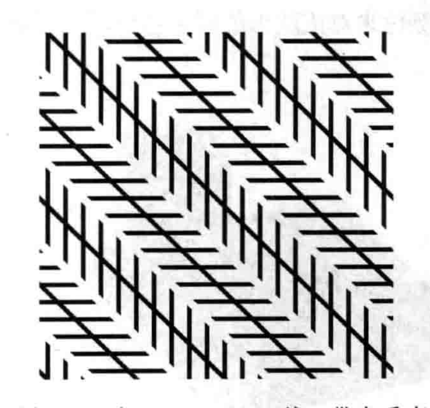


图 2.22 左氏（Zollner）线。带有垂直或水平短标记线的交叉对角影线，会使得我们对它们的视觉感知朝相反的方向旋转。事实上，这些对角线是完全平行的

人类并不使用大脑中的量角器来测量角度，但能对它们进行比较并注意到差异。对于跨阶的亮度变化，会放大角度上的差异来检测边界并增大感知的变化。在腕表表盘上没有任何标记的情况下，我们到多能做的是说出最接近分钟的时间。一分钟对应分针 6° 的运动。但如果存在两组只有几度差异的线条，那么我们会注意到这一差异（通常将它判断得更大）。

关于角度的抑制，意味着有垂直或水平短标记线的垂直交叉对角线可以改变我们对主线方向的感知。如图 2.22 所示，这会使得很难或不可能正确地比较对角线的角度（它们实际上是平行的）。

2.7 局部到全局层次结构

场景中元素的解译，很大程度上依赖于将它们组合在一起形成结构和对象。一个非常简单的例子（见图 2.23）只使用点显示眼睛通过连接最近的邻点来构造对齐。点和线的类似组合用于将定义不明确的特征边界连接在一起，然后为视觉解译创建对象的轮廓（见图 2.24）。这种组合的一种简单扩展是，包含随时间流逝的最近邻点来产生《星际迷航》中的星空运动。下面详细讨论时间上的比较。

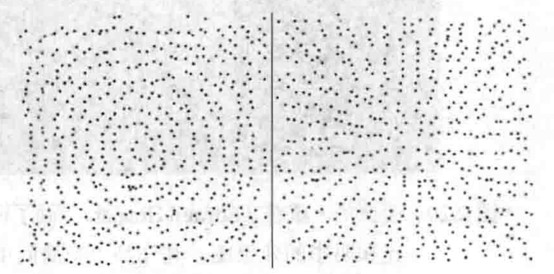


图 2.23 连接每个点与其最近的邻点时，会产生放射状或环向对齐的印象

为使抑制工作更为有效，组合是必要的。环绕、对比帧所致亮度变化的常见错觉，取决于使用中心区域进行组合的帧。在图 2.25 所示的例子中，将帧从中间分开的一条黑线改变了图像的外观，并减少或消除了对灰色中心区域亮度的感知差异。若没有这条线，亮帧和暗帧会与中心组合，而抑制改变了所表现出的亮度，使周围较暗的区域看上去更亮，反之亦然。

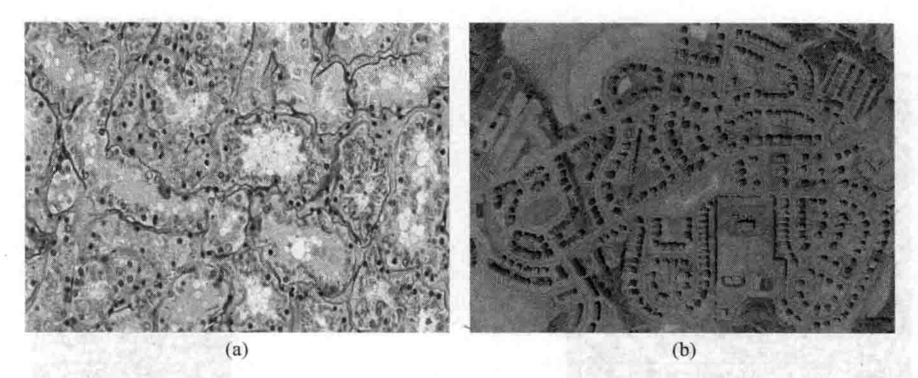


图 2.24 线和点的视觉组合：(a)生物组织切片的细胞边界是通过对断开的染色区域进行组合所构成的；(b)因为有房屋来界定，所以航拍照片上的街道非常容易检测

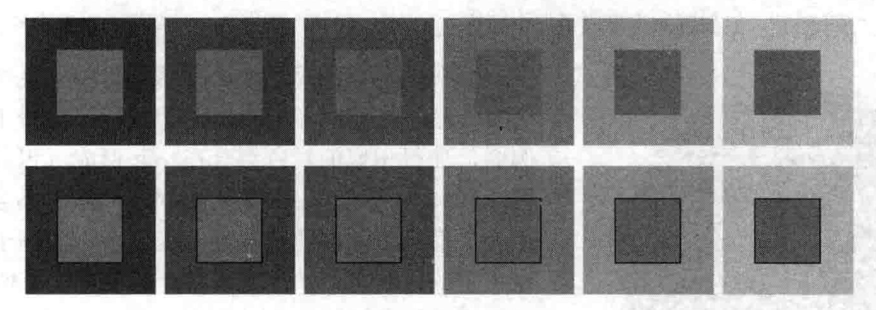


图 2.25 在顶行，亮和暗的环绕影响我们对（相同的）中央灰色区域的视觉比较。在底行，这种效果被分隔边界降低

彩色区域组合如图 2.26 所示。根据邻近的是蓝色还是黄色，红色条纹的视觉解译是不同的。组合的另一个结果是，属性（例如颜色）会延伸到周围区域。在图 2.27 中，线条之间的白色区域是相同的，但并未感知成那样，因为最近的线的颜色，或橙色或品红，和它们组合在一起，因此视觉上它们遍布于这些区域中。

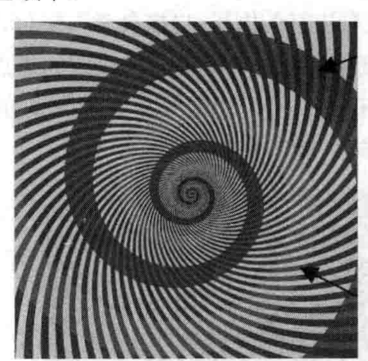


图 2.26 颜色解译受周围组合的影响。两个螺旋中标记的红色实际上是相同的，但因为分别邻近蓝色和黄色条纹，所以被感知为品红色和橙色

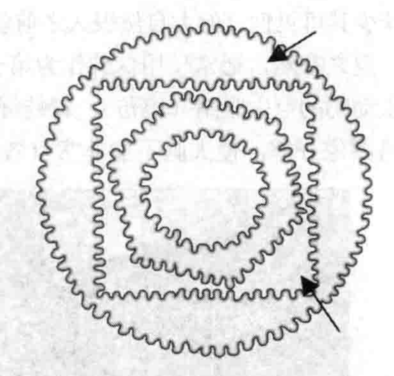


图 2.27 箭头所标记的区域是相同的白色，但来自相邻线条的颜色通常会被视觉延伸到这些空间，并被感知为对这些区域添加了颜色

组合与抑制的结合会产生如图 2.28 所示的错觉。有角度的阴影导致线条看上去稍微偏向了相反方向，而线条的不同部分通过组合连接起来，所以整个图被感知为一个螺旋。但实际上，该图案由圆圈组成，这可以通过跟踪其中的一个来验证。这是一个空间组合的例子，但时间分组以大致相同的方式工作。旋转的螺旋（见图 2.29）看起来通常会产生无尽的运动。

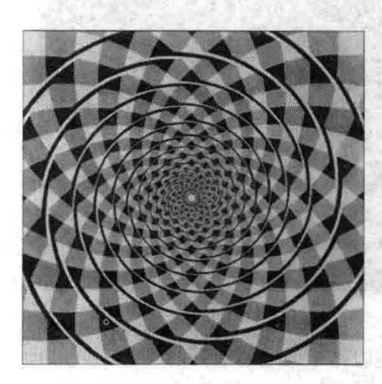


图 2.28　弗雷泽（Fraser）螺旋。圆环视觉上看起来是“平铺的”，但被感知为形成了一个螺旋

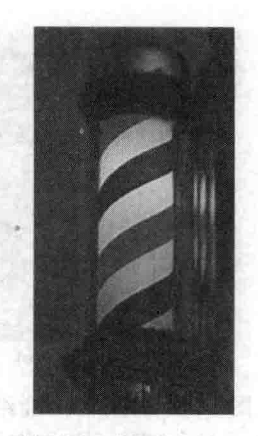


图 2.29　转动的螺旋形或螺旋图案，如理发店招牌产生一种背景垂直于线条的运动的错觉

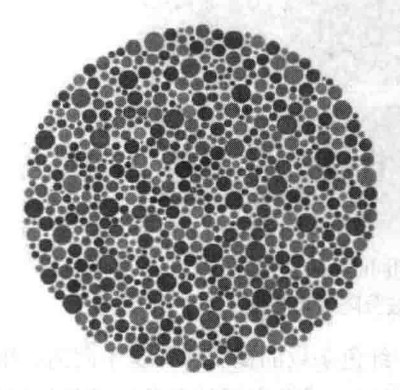


图 2.30　石原的色盲测试图之一

常见的色盲测试的核心是将图像中颜色相同的特征组合在一起。在石原（Ishihara）测试中，安排了一组彩色的圆圈，使得相似的颜色组合在一起形成可识别的数字（见图 2.30）。对不会分辨这些颜色的人而言，不会出现同样的组合，而且对数字的解释会改变。所示的例子是一组对红绿色盲的诊断之一，这是最常见的色盲，男性中有近 1/10 的人患有这种色盲。在该例中，正常色觉的人应读出数字 74。红绿色盲会将 4 视为 1。全色盲的人看不到任何数字。但色盲的人通常能看透迷彩，这要优于视觉正常者。

通过干扰人们将图像各部分组合在一起的视觉能力，迷彩被用来隐藏对象。在第一次世界大战前，军方发现了这种现象，当时美国海军舰艇传统的白色（西奥多·罗斯福的“白色舰队”）被不规则的灰色和蓝色图案所代替，以减少其可见性。但大自然很久之前就明白这个道理，而且有丰富的例子。“捕食者”要对其猎物隐藏自己，反之亦然，通常使用伪装作为第一道防线。当然还有其他的视觉可能性——蝴蝶的斑点看起来就像更大动物的巨大眼睛（模仿），青蛙鲜艳明亮的颜色警告其有毒——但通常（见图 2.31）的目的是隐藏。将图像分解，使大脑不能非常有效地把各部分组合在一起，就防止了识别。

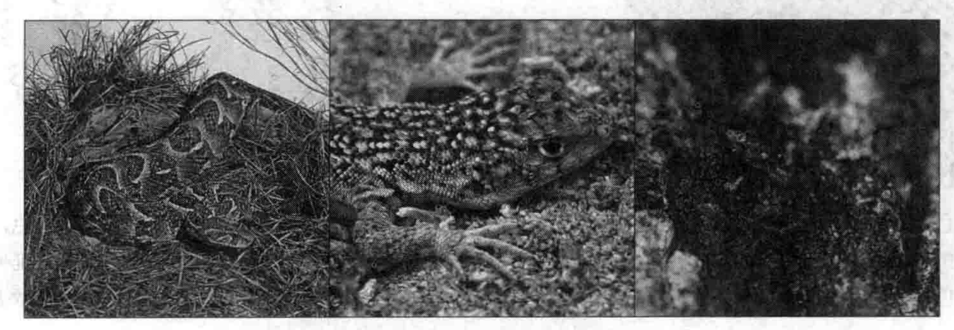


图 2.31　天然伪装（蛇、蜥蜴和青蛙）

运动可以破坏由迷彩导致的错觉，因为移动的物体（或其部分）吸引注意，而且如果有图像的多个部分被观察到以协调的方式移动时，它们就会被组合在一起，这样隐藏的对象就会出现。人类的视觉试图将移动的特征或点视为以协调的方式移动的刚体及很容易组合的部分。改变光照或彩色滤镜来

观察场景和图像，也往往会揭示对象。但在自然界中，保持静止和依靠伪装的能力是一个高度演化的策略。

组合可在许多空间（和时间）尺度上进行。在典型的场景中，可能是线、边缘和其他特征组合在一起被感知为一个物体，但接下来该物体会与其他特征组合，形成更高层次的组织，以此类推。分组层次违规时，会导致必须由大脑来解决的冲突。有时，这可通过只看一种解释而忽略其他解释来解决，而有时大脑必须在解释之间来回切换。图 2.32 显示了两个例子。如果红色的物体被感知为前景，它就被视为一个苹果核。如果苹果周围的区域变成前景，那么它会作为两个相向的人的侧面像出现。该图可被视为一个年轻的姑娘或一个老妇。对于所有此类图像，即有多个解释或前景背景反转产生不同认知的图像，有些人最初只能看到一个人，或存在其他可能性，或在识别两者之一时存在困难。但一旦设法“看出”了两种解释，那么一次看到两者是不可能的，而且对观看图像的大多数人而言，这两种可能性每几秒钟交替一次。

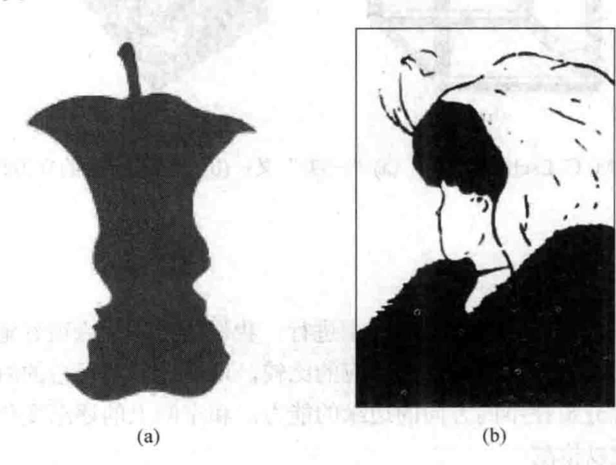


图 2.32　两种不同解译的错觉：(a)两个相向的侧面像或一个苹果核（慈善机构 Share Our Strength 抚养饥饿儿童的符号）；(b)年轻的姑娘或老妇

这些例子为我们了解大脑如何处理图像提供了令人感兴趣的展示，但对依赖于对图像进行视觉检测来获取物体信息的那些人提出了警告。要看到过去的伪装，连接对象的不同部分并识别其存在，我们必须拥有一个对象是什么的较好存储模型。但当我们以一个强大的模型趋近图像并试图将图像的各部分连接到模型时，就必须忽略其他的解译。图像中不符合该模型的任何内容都可能被忽略。

在 Tony Hillerman[⑤]的一个侦探故事中，春纳瓦霍（Navajo）[⑥]侦探 Joe Leaphorn 解释他如何寻找痕迹。FBI（联邦调查局）的人问：“你在找什么？” Leaphorn 答道：“没什么特别的。你不是真的在寻找什么特别的东西。如果你这样做，你将看不到你不在找的东西。”但学习如何寻找一切和什么都不找是一个很难掌握的技能。我们中的大多数人仅仅看到我们希望看到的东西，或至少是我们之前就已获取标签和模型的熟悉的东西。

艺术家埃舍尔（M. C. Escher）[⑦]创作了很多组合层次局部一致但全局却不一致的作品，来创造艺术的冲突和对抗效果。图 2.33 显示了他创作的一些谜图。在第一幅图中，叉端表示一种边缘的线，在叉底则变成了不同的东西，沿途交换了内侧与外侧。在第二幅图中，边缘的前后顺序改变了。挡住其他边缘的边缘，通过与放置在这些边缘后面的角组合，被感知为前面的边缘，反之亦然。无尽的楼梯

⑤ Tony Hillerman（1925 年 5 月 27 日—2008 年 10 月 26 日），美国侦探小说家。——译者注
⑥ 美国最大的印第安部落。——译者注
⑦ 摩里茨・科奈里斯・埃舍尔（M. C. Escher，1898—1972），荷兰艺术家，专门从事木版画和平版画的创作。——译者注

使用透视变形来形成一个没有顶或底的封闭路径。在所有这些情况中，信息的局部解译是一致的，但不一致的非全局分辨是可能的。

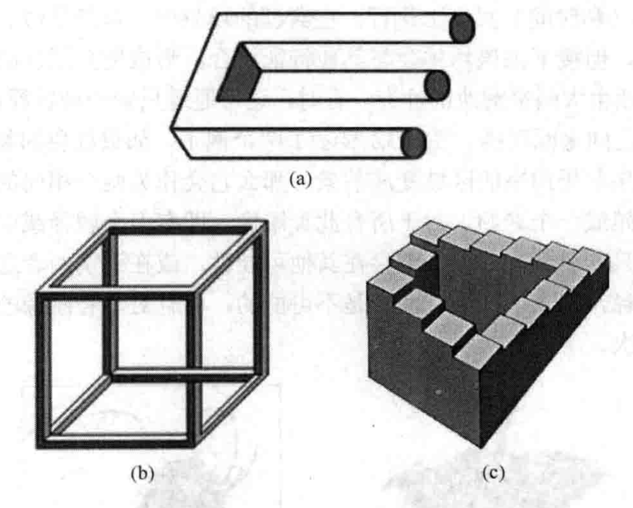


图 2.33　素描（由 M. C. Escher 创作）：(a)“三头”叉；(b)一种不可能的立方体；(c)无尽的台阶

2.8　时间问题

比较和抑制既可以在时间上也可以在空间上进行。我们的视觉边缘很好地与检测运动关联到了一起。某个时刻与很短时间之前的一个时刻的响应的比较，可通过一个很短的时延完成。稍微复杂的连接检测边缘的运动，具有分辨在不同方向的边缘的能力。和空间上的逐渐变化一样，亮度上的逐渐变化和慢速运动被忽略且难以检测。

时间抑制与适应或缺失并不是同一回事。我们的眼睛需要花一定的时间去适应场景亮度的变化。其中一部分是虹膜开闭瞳孔的响应，让更多或更少的光进入眼睛。适应响应部分是细胞内的化学变化，以便放大微弱的信号。前者需要数秒时间，而后者需要数分钟时间。

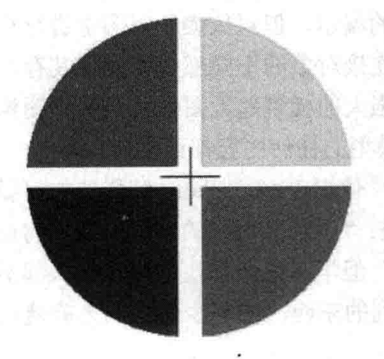


图 2.34　说明余像的靶。凝视中间的十字大约 1 分钟，然后转移目光再看一张白纸，将会看到与原图颜色互补的图像

眼睛短暂地注视一幅固定的图案或一个彩色目标，化学上会使得视杆和视锥缺失。之后盯着一个空白页，将会产生原图像的负片或相反的图像。图 2.34 显示了一个简单的例子。一动不动地盯着圆圈的中心约 60 秒，然后将视线移开。因为对颜色敏感的视锥已缺失，圆圈的余像将由相反的颜色（绿色对应于红色，黄色对应于蓝色等）组成。原始图像越明亮，余像就越强。

运动感知十分重要。它提醒我们可能出现威胁或机会的周围环境的变化。推断运动的能力位于使得我们实现诸如抓住一个抛出的球这类运动的能力的中心。跟踪和推断为我们注意运动的不连续性提供了二阶机会。如果某个运动着的特征突然改变速度或方向，则它也将会被注意到。这是因为跟踪运动和注意到数据图表中细微变化的能力非常有用，也是原始数据的导数图通常能在视觉上揭示信息的原因。

若图像序列以非常慢的速率出现，以至于可以视为单幅图片的一个序列，或非常快展现出连续运动

的错觉，则其解译会完全不同。摄影师埃德沃德·迈布里奇（Eadweard Muybridge）的很多电影和电视作品毁誉参半。他曾参与了 Leland Stanford 邀请的一场赌局，即确定一匹飞驰的马是否会出现四脚同时腾空的情况。迈布里奇安装了一排照相机，相机配有绊网，以便在马匹跃过时对其进行摄影，生成一组相片。单独地观察这些相片足以确定赌局的胜负。但迈布里奇发现快速地浏览这组图片将会产生连续运动的视觉假象。电影工业几乎同时诞生。图 2.35 显示了由迈布里奇拍摄的 12 张奔马图像的一个序列。

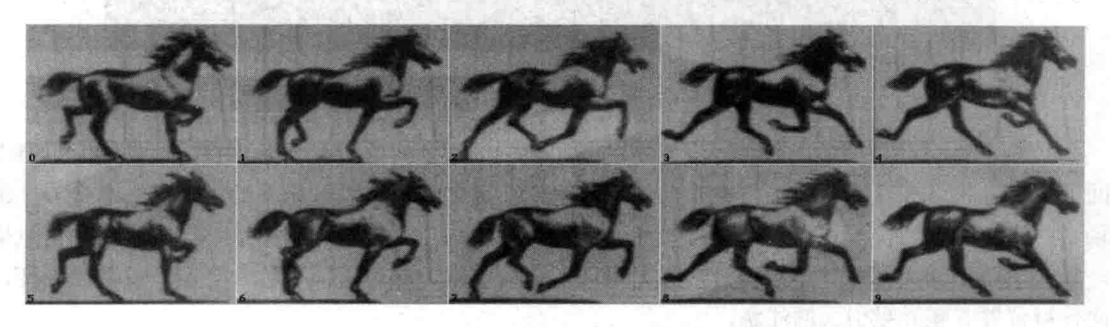


图 2.35　迈布里奇拍摄的奔马图像的一个序列

在观察一组图像时，可能会出现混叠现象。通常，人们假定一帧中的一个特征对应于下一帧中的某个特征，如果它们在颜色和形状上几乎一样，而且它们靠得非常近（即组合）。我们所熟悉的转轮辐条反方向转动就是视觉错误的一个例子，它出现的原因是一帧中的一个特征与下一帧中的另一个特征进行了匹配。通过观察这一现象，就出现了频闪成像的一小步，在频闪成像中，一组图片以与某些现象重复出现的速度匹配或近似匹配的时间间隔来拍照。这通常被用来研究旋转或振荡物体、下落的水滴等。即使齿轮上有不同的齿或以不同的水滴作为主题，每幅图像也都显示了几乎一样的东西。我们在视觉上假定它们是一样的，感觉不到差异。稍微改变图像的定时（通常通过控制光源闪烁的速率），可以从更快的重复中提取出所发生的细微图案和运动。

如果图像序列慢于 15 帧每秒，将不会感觉为连续的。事实上，10～12 次每秒的闪烁图像，会让人有如同癫痫发作时的痛苦。若以更高的速率运动，眼睛和视觉系统的暂态响应就会看到连续的动作。电影通常以 24 帧每秒记录，而电视广播则以 25 帧每秒（欧洲）或 30 帧每秒（美国）记录。在所有这些情况中，序列图像被解译为连续的。大多数正在使用的视频监视系统存在的问题之一是，图像不是以 30 帧每秒而是以每隔几秒存储的单个帧（或者通常以只有一半的垂直分辨率的单个帧）记录的。结果是一组缺失了一些重要线索的静止图像。我们不仅通过静止图像，还通过人如何移动来识别人，而在录制中缺失运动轨迹，因此观察到姿势而非运动是可能的。

正是由于人类视觉在一个相对很窄的范围能够很好地检测物体或亮度和颜色的变化，所以能很好地处理在相对较窄的时间范围内发生的事件。不依赖于诸如高速摄影机这样的设备来捕获，人眼是无法看到持续时间非常短的事件的（见图 2.36）。高速成像和频闪成像技术都是由麻省理工学院的 Harold.“Doc” Edgerton 在 20 世纪三四十年代开发的。

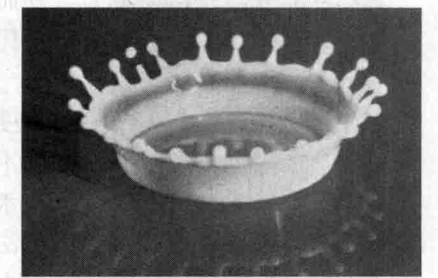


图 2.36　Doc Edgerton 的著名高速摄影作品：溅起的牛奶滴

同样，需要很长时间（数分钟）才能发生的事件，视觉上检测其变化并不容易。即使是图 2.37 中放在一起的钟面图像也没有提示出所有的不同，且当其中一幅图像必须从记忆中回忆时，结果将更差。在计算机中捕获图像并且计算差值清楚地显示了分针的运动甚至时针非常微小的转动。它们的不同可以清晰地看到分针的运动，甚至变化得更为缓慢的时针的运动。延时摄影是用于突破时间界限并使得慢速过程的视觉研究成为可能的重要工具。

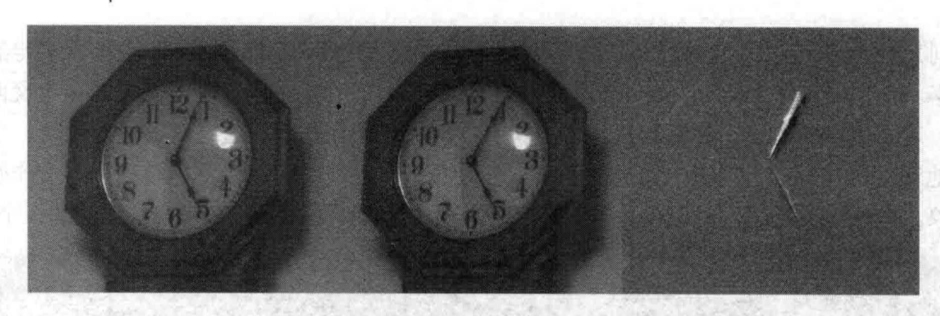


图 2.37　相隔大致 1 分钟拍摄的办公室墙上的时钟图片及它们之间的差

考虑到视觉对运动的敏感性，令人吃惊的是，我们的眼球实际上几乎是一直运动着的。只有在中央凹处，高分辨率观看才有可能，在眼窝中移动眼球，可使得这个高分辨率的小区域一个接一个地观看场景中许多的单独位置。浏览场景中看起来令人感兴趣的点，可将大脑用于解译和判断的信息聚集在一起。场景中的大部分并不被检查，除非出现了边缘、线、颜色，或颜色、亮度、纹理的突变，或方向使得位置有趣并吸引人的注意。

在某种程度上，随着眼球的移动及每隔几百毫秒不同的图像落入视网膜，大脑将辨别它们，并将这些信息插入到大脑中构建的感知场景中。尽管整幅图像在视网膜上不断地移动，但它只是在被注意到的心理场景中的相对运动。眼球的这种运动也用于消除盲点，盲点是视网膜上视神经进行连接的位置，该处没有感光器。在视场中缓慢移动光源，同时使眼睛盯住一点不动的实验，很容易验证盲点是存在的，但因为眼睛运动收集了足够的信息来填补我们所解译的场景的感知模型，故在实际观看时从不会注意到盲点。但在实际场景中有可能有大量的信息未被注意到，很显然，这会存在大量的解释（在章鱼中，大自然已经重新排列了眼睛的结构，使得视神经连接到视网膜的背面，因而不存在盲点）。

一个场景的感知图像与同一场景的照相记录很少相同。后者以一种永久记录存在，并可以在以后对每部分进行详尽的研究。而场景的精神意象是短暂的，且其中很多都被来自我们视觉边缘或来自其他相似场景的记忆中的低分辨率信息填充。科学家们需要记录图像而不只是观看它们，直到很久后发现其中的重要特征。

处理原始数据的眼睛和大脑连接的结构与功能的进化，一直响应于环境和生存的挑战。如前所述，不同的动物显然有不同的需要，而这导致了不同类型的眼睛，以及不同类型的处理。很明显，苍蝇眼睛传送到大脑的少许特定信息，足以引发合适且成功的响应（例如，接近表面时引发着陆反射，苍蝇调整自己的身体以便让脚先接触表面）。但是对于绝大多数“高级”动物而言，通过视觉系统收集到的信息类型和自动应用的解释则非常多样化。

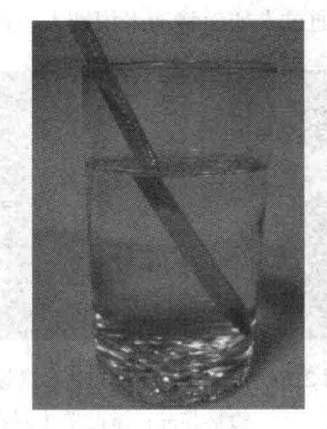


图 2.38　铅笔的明显弯曲及其放大是常见的视觉效应

例如，当人们看放在一玻璃杯水中的铅笔时，铅笔看上去以一个角度弯曲（见图 2.38）。理性上我们知道，这是由于水与空气的折射率不同所致，但呈现给我们大脑的视图仍然存在弯曲。终身都用矛在浅水区捕鱼的南太平洋岛上的居民学会了如何自动补偿这一弯曲的方法，但他们的眼睛呈现给大脑的图像仍然包含这一光学弯曲。有证据表明，像鹭这样捕鱼的鸟类对其视觉系统有一个内置的用于偏移的校正，而且从所感知图像的观点来说，它们直接扑向鱼以便抓住它。它们“看到”一幅校正后的图像，而不是我们观察到的物理表示。相反，存在某些种类的鱼，它们看到水面上的虫子后，会向它们吐液体将其击落，从而把它们作为食物。这也需要对光学位移的类似校正。

大脑中的计算可以以与校正诸如光在水中的折射率这样一种纯粹的光学效应完全相同的方式对一

幅图像进行扭曲，这似乎很奇怪。但这只是与已描述的调整在程度上的不同，在已描述的调整中，人类视觉会校正表面的阴影及随亮度和颜色变化的光照。为创建连贯且准确的或至少足够准确的场景描述，该过程会将原始视觉数据与大量“我们已知的世界内容”组合在一起。

人类擅长于跟踪运动的目标并预测其路径，并会考虑空气阻力和重力（例如，棒球外场手会抓住正在飞的球）。大多数的动物并不具备这样的能力，但少数学会了。我的狗总能朝着抛出的球的当前位置追逐它，这就形成了数学上的抛物线。的确如此，这并不需要太多计算，但并不是最优的。另一方面，我的猫沿直线跑向球即将要到达的地方。它求解了数学上的抛物线，它猛扑向球，但并没有像狗一样把球带回给我。

我们常说的“手眼”协调涉及视觉场景中所发生事情的大量细微计算。能够很好地跟踪超过 100 英里每小时运动的球并以受控方式将其击中的职业棒球或网球运动员，具有很强的条件反射，但这是以大量的视觉敏度和处理为前提的。在用来训练年轻棒球运动员的一种练习中，设置有一个自动投球机将球射穿过本垒板，而处在击球区的运动员必须及时地识别出染有不同颜色斑点的球并做出反应。从这样的练习中，他们学会了认出以快速直球、曲线球及滑球投出的球的接缝的不同旋转。由于球从投手到接球手总共飞行时间大约是 2 秒，因此还需要一定的时间激活肌肉来做出反应，这就需要极快的视觉处理。如前面所提到的，人眼有大约超过 1.5 亿个感光器，而每个感光器有 25000～50000 个处理神经元以并行方式工作，以便提取可使我们更好地生存的信息。

2.9　第三维

多数人都有两只眼睛，且至少具有由两只眼睛的立体视觉以某种方式为我们提供所见物体的距离信息的基本想法。很多动物的眼睛位于头的两侧，两个视野的重叠很小，且并不强烈依赖于立体视觉。大多数肉食动物和在树上跳跃生活的动物的眼睛是朝向前方的，这表明立体视觉是很重要的，但立体视觉不是确定距离的唯一方式，也不是具体的定量工具。

人类通过在眼眶中旋转眼球得到立体视觉，把同一特征带到每个眼睛的中央凹处（就像视觉皮层里发生匹配时所做的判断一样）。依赖于注意力从一个特征移向另一个时，眼睛是否需要向内或向外旋转，肌肉到大脑的反馈告诉我们一个特征是否比另一个更近。注意，这并非一个物体比另一个物体近多少或远多少的测量，它只适用于两个点的比较。只有通过扫视场景四周并建立大量的两点比较，才能使大脑构建出许多未知的相对距离图。第 1 章介绍了立体距离测量的基本三角方程。匹配场景之间的点并建立一个相似的距离图的计算机算法，将在第 14 章和第 15 章中介绍。

如果一幅图像中有几个相似的特征，那么立体视觉可能会变得混乱，因此会有多个匹配（对应于不同的表观距离）。在自然场景中很少出现这种情形，但在对重复结构的成像中则会产生问题。

关于立体观测（匹配来自左眼和右眼视图的细节进而融合立体像对的能力）的一个令人陶醉的发现是随机点立体图（见图 2.39）。Bela Julesz 于 1959 年在贝尔实验室证明视觉系统可以匹配对单只眼睛看上去混乱且没有结构的点图案来形成立体图像。点的轻微横向偏移被解释为视差并产生了深度信息。

通过移动单个眼球所产生的序列图像也能产生立体的深度信息。就像头从一边移动到另一边一样，场景中特征的相对侧向运动是与距离成反比的。有一种理论认为，眼睛位置并不能很好地适合于立体视觉的蛇通过将其头部从一侧移动到另一侧，来更好地对距离进行三角测量而发动攻击。

立体视觉只对非常近的物体有效。在 100 英尺之外时，角度差异变得太小以至于难以发现。此外，还有大量的人因为这样或那样的原因而不具备有立体视觉能力（如这是童年弱视的典型后果），但他们仍然在三维世界里表现得不错，如可以驾车、玩高尔夫球等。图像中还有其他线索用来判断距离。

如果一个物体遮掩了另一个物体的一部分，那么它肯定离我们的眼睛更近。优先性似乎是判断物

体距离顺序的绝对方法，至少是对那些沿着同一视线的物体。但是，也有关于识别和物体的简单形状的内在假设，如图 2.40 中的例子所示，其破坏产生了错误的解译。

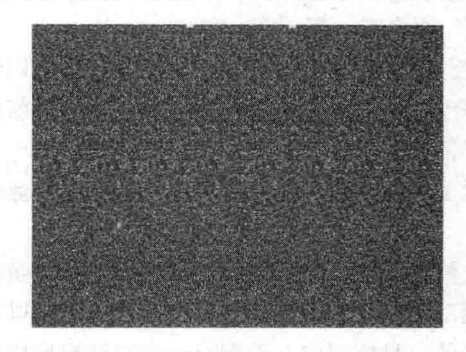


图 2.39　随机点立体图展现了在平面上的一种圆环（凝视图像，眼睛会提取匹配的图案，然后将其融合成一幅具有深度的图像）

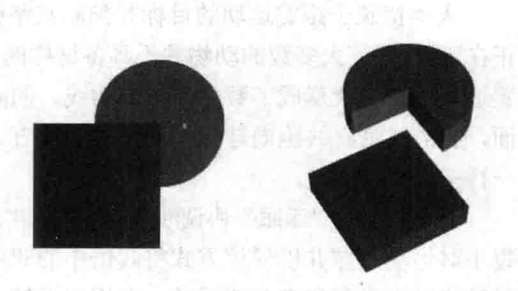
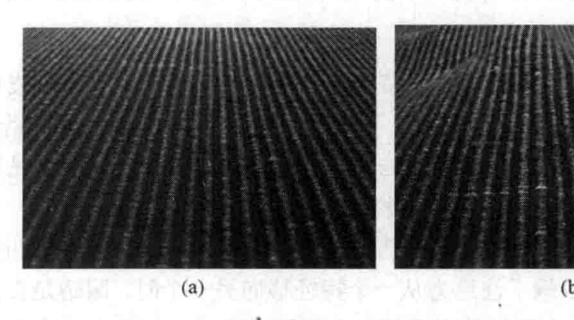


图 2.40　很明显，蓝色方块在红色圆圈的前面，但它也可能不是一个圆圈。另一个角度的视图显示红色物体实际上在蓝色物体的前面

相对尺寸在判断距离时也起着重要的作用。在熟悉的场景中，所识别的许多物体具有落入很窄范围的尺寸——最典型的是高度。与此密切相关的是对透视规则的理解——平行线在远处看起来会会聚于一点（见图 2.41）。出现此类线的视觉解释通常是，对应于从观察者向远方延伸的平面，而直线的不规则可解释为表示感知表面上的隆起和下凹。驾车通过乡村时，观察犁过的田地便可以提供一个简单的例子。



(a)　(b)

图 2.41　会聚的线被解释为按照透视规则会聚的平行线，所以将感知的平面表示为远离观察者。直线表示平坦的表面(a)，而不规则可以解释为表示感知表面的隆起和下凹(b)

通过比较视野中特征的表象尺寸，我们可以判断物体及与之相距很近的其他物体的相对距离。同样，基本假设对成功很重要，且违反特征的直线排列或尺寸的不变性，将会产生错误的解译（见图 2.42）。

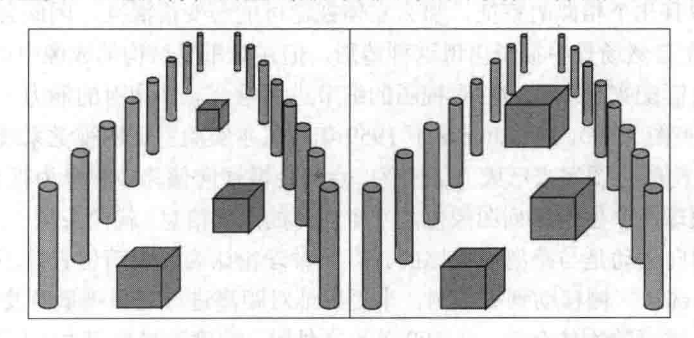


图 2.42　在这些图示中，通过假定粉色的标杆大小相同且以笔直的平行线排列来建立距离的期望值，会聚的平行线是透视的结果。在左图中，按照这一解释，绿盒子看上去是不变的。但在右图中则不然，要么认为盒子的大小不等，要么认为标杆不符合我们的期望

主要特征尺寸不变这一假设的一个简单推理是，可以使用表面上标记或特征的尺寸来估计距离或角度。假设表面上所呈现的标记或纹理是随机的、各向同性的，且表象尺寸或横纵比的视觉变化表明了距离或方向上的不同，这是符合逻辑且通常是正确的（见图 2.43）。违反基本假设同样会导致错误的结论。



图 2.43　沙滩上的石头。石头大小相同且是圆形的，假设给出了视角以及到更远位置的距离

真实场景中可能会出现其他线索，尽管它们与观看近的物体或来自显微镜或实验室的图像的联系并不大。例如，大气霾雾会让距离远的物体看上去更蓝（若霾雾为烟雾，则为褐色）和更模糊。文艺复兴时期的画家们对这些线索掌握得非常好，他们在场景中使用正确的几何透视来表示大气霾雾。但由已知的几何关系到理想的表示，完全不同于从其成分部分已知的场景中提取几何信息。

2.10　工作原理与工作内容

针对眼睛及视觉通路中其他部分运作的算法，人们提出了几个非常可靠的模型。视网膜神经元的最初几层以可以解释局部和时间抑制的机制的方式相连接，视觉皮层中右眼和左眼信息的交织和用于立体观测的两幅图像的融合相一致。这样的模型有几个目的——能够被生理探针和外部刺激物所测试，而且它们形成了试图从图像中提取相同信息的计算机技术的基础。即使其结果作为神经元的功能的实际描述是失败的，后一目的仍然能达到。但当它们可以有效地描述视觉系统至少某些部分的工作原理时，因为其只工作于图像的零碎层，而不工作于完全形态层或信息层，因此并未告诉我们关于所见内容的太多信息。

几年前，我是一宗刑事审判聘请的专家鉴定证人。手边的问题是判断在谋杀现场的监控录像带是否对鉴定疑犯有价值。录像中的图像被彻底使用电脑增强并呈现给陪审团，陪审团的结论是：从这些图像中不能辨认出被告当前的外貌。这些画面在空间和色调分辨率上实际上都相当贫乏，甚至不能将它们与谋杀发生所在城市的大部分人群区分开。辩方的工作是提醒陪审团图片中是否包含有足够的信息来辨认被告。在该案件中，非常有趣的是，许多目击者中没有一个人可以从一排人中指认出被告。人眼与那些廉价的黑白监控摄像机相比，无疑会有更高的分辨率且更敏感，但由于各种原因，在场的人们无法辨认出罪犯。

在那场审判中，我被法院接纳为计算机图像处理领域（解释图像处理过程）和人类感知领域（解释在场的人可能会看到什么，可能看不到什么）的专家。关键是我的学位和背景并不是生理学领域，而是物理学和工程领域，怎么能够作为人类视觉处理的专家呢？我所做（并被法院接受）的回应是：在视紫红质或神经元层，人类的视觉系统是如何工作的问题并不重要，重要的是人类视觉能够从场景中提取什么信息。我当然理解从图像中可以看到什么，因为我已经花了 50 年试图去找到计算机提取某些相同信息的方式（大多数情况下使用非常不同的算法）。顺便说一句，要实现这一目标，可能还需要毕生的努力。

人类视觉“如何工作”和“看到什么”之间经常会有混淆，且通常会被“为什么”的问题搞得更为复杂。在描述针对图像分析的某些方面的计算机算法时，解释信息提取的步骤（“如何”）与计算出的结果（“什么”）关系密切。但该算法可能不是（事实上通常也不是）获取信息的唯一方法。图像分析中很多的重要步骤有几种可以提取相同结果的等价方法，而且其中的每一种通常可以不同的方式编程来利用不同计算机结构的独特性。当然，这些实现都与大脑中的神经元所进行的处理一致。

David Marr 在其“Vision”一书（Freeman，1982）中非常有说服力地指出，混淆于“如何”和“什

么”让包括他在内的许多研究者在追求（包括人类和动物的）视觉理解时陷入了死胡同。将视觉皮质中不同部分内的微小电信号映射为呈现给眼睛的刺激函数，或者测量单个视杆和视锥细胞的谱响应，必然是最终理解视觉系统如何工作的一个重要部分。它可以实现，至少是部分可以实现的实验，但不清楚的是它到底能告诉我们“什么”。另一方面，蛙眼对于一个很小的黑色移动物体的反应测试，会引起关于“为什么”的推测（如检测一只正在移动的昆虫：食物）。



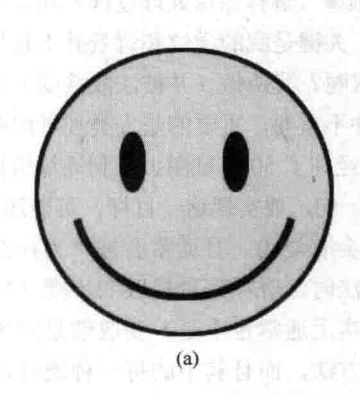
图 2.44　理查德·尼克松的滑雪鼻子、深色的眉毛和阴暗的眼睛、向后的发际线以及 5 点钟阴影的下巴被漫画家用于创建一幅即刻就可以识别的讽刺画

研究人员已进行了许多实验来确定人们会看到什么，通常包括在受控的环境中展示人工刺激并在引入小变化时对响应进行比较。这已产生了一些有用和让人感兴趣的结果，但还不足以阐述场景中视觉解释的问题。关键并不仅仅是人们可以检测（“看”）到某种刺激，而是可以解释在复杂场景中该刺激的含义。当将它们组合为一种语义表示时，用“理解”替换“看”来强调图像中的个体线索，仅适用于理解的时间。换句话说，必须将图片转换成等价的“文字说明”。研究人类在看所有图像时所犯的错误通常更具有揭示性，这包括但不限于各种视错觉。

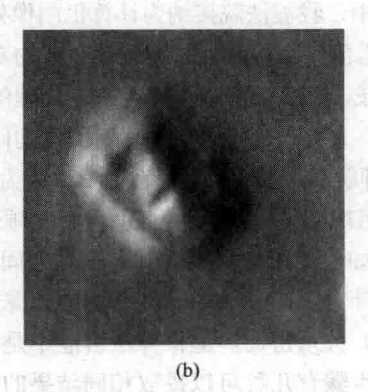
艺术家在代表性的画作甚至漫画中描绘了（或忽视了）重要的线索，他们在一幅讽刺画中会夸大某些所选的特征，如社论漫画家将熟悉的政治人物创作成非常歪曲却又能够容易识别出的形象。对于很多人而言，这样的漫画可能比一张实际照片更能表示一个人的更多事实（见图 2.44）。

2.11　看到不存在的东西，反之亦然

困扰目击证人证词和鉴定的一个问题是，人们倾向于看到（即从场景中选出）熟悉（即已经心理认可）的东西。就像青蛙认出昆虫的能力一样，固化在我们大脑中的一种本能是找到脸。婴儿自出生就会识别和跟踪脸。我们太擅长于这一点，以致即使只有很少的线索（如两只眼睛、一张嘴），也能看出一张脸，而不管其是真还是假。无处不在的“笑脸”漫画有足够的信息而被识别为一张脸。在合适的光照下，以一定的方式观看时，火星上的一座山也被识别为一张脸（见图 2.45）。



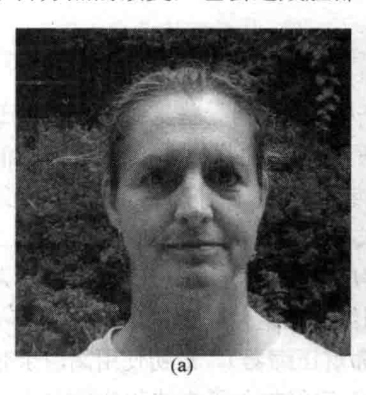
(a)



(b)

图 2.45　只需很少的线索就能触发对脸部的识别：(a)无处不在的笑脸；(b)视角和光照合适时才出现的“火星上的脸”

但要识别一张特定的脸，如识别为奶奶，则需要有更多的线索。实现面部识别的计算机程序使用维度比，例如眼睛间的距离和两耳末端间的距离之比，或者口到下巴的距离和鼻尖到下巴的距离之比。这种比值的优点是对图像的大小、方向或视点在相当大程度上不敏感。但这是一个算法，所以强调的是“怎么做”，而不是“做什么”。人类面部识别似乎使用了更多的或不同的线索，但毫无疑问，对脸部比例即使只做几个百分点的改变，也会造成脸部不能识别（见图 2.46）。


(a)


(b)

图 2.46 改变维度比（如眼睛间、耳朵间的水平距离，嘴巴的宽度等，或垂直尺寸，如鼻子的长度、嘴巴到下巴的距离、前额的高度等）强烈地影响着识别脸的能力

警方通常会由目击者的描述画出草图。如图 2.3 所示，在嫌犯抓获后，将这些草图与嫌犯的真实照片进行比较表明，只有少数的脸部特征可能会被注意到并变换成心理上的讽刺画而非实际表示。目击者和嫌犯的种族差异，使得挑选出可能识别该人的那些特征特别困难。我们找出了对辨别熟悉的种族最有用的特定细节，而这些对其他种族而言却不是非常有帮助（“他们看起来都很相像”）。

最后，当某人或某物在一幅照片中被（正确地或错误地）识别出来之后，我们会用记忆中的一组特征赋予该人或该物的语义表示。一个典型的例子可能是从侧面看朋友的脸，但“知道”在另一侧脸颊上有一颗痣，并相信我们此时也看到了这颗痣。这就导致了大量的目击者问题。如果一位目击者认为他已经识别出了某人或某事，那么他通常会充满信心且诚实地作证说看到了实际上并没有发生的事情。这可能包括非常具体的内容，如衣服、文身等。在肇事逃逸案中，某目击者确信一辆车尾保险杠上有一张特殊的贴纸，其实当时她处在不能看到车尾的位置上，只不过是她对同样型号、同样颜色的车很熟悉，而那辆车确实有这样一张保险杠贴纸。

我们会毫无意识地这么做。当你经过邻居的房子时，如果瞥见某人在草坪割草，而你“期望”割草的人是处于青少年的儿子，那么你很有可能会“看到”他的外貌、发型、衣着等细节，这些可能看不到，甚至可能不存在。通常，这个过程是有用的，因为它给我们一种位置感，而且这种感觉在不需要额外的时间和精力的情况下通常是正确的。这样做会犯一些错误，但幸运的是，这些错误通常并不严重，且造成的结果不过是瞬时的尴尬。几十年前，当我大儿子还是学龄前儿童时，我剃掉了从他出生之前一直留着的胡子。头两天他都没注意到，直到向他指出来（而之后他对这一变化感到不舒服）。

通常我们希望出现我们所“知道的”东西，或至少期望这些东西出现。与我合作讲授图像分析课程多年的一位同事有一张他特别喜欢的图片，图片中她抱着一个非常可爱的宠物，但宠物现在已经去世了。很不幸，这条狗是黑色的，她也穿着一身黑色的毛衣，因而冲洗出来的照片非常暗（不存在更大动态范围的底片）。多年来，她一直要求我和我的学生处理这幅图像，以便能看到图片中的那条狗，但我们失败了，因为从像素值来说，图像中没有任何东西——它们的明暗度全都接近黑色。我的一名学生拿到该图片的一份扫描件后，煞费苦心地画出了这条狗的可信轮廓。但她的第一反应是“那条狗不对！”她的大脑中存储了一幅特别的图像，其中包含了其他观看者无法看到的信息，而她相信自己可

以看到那些信息。显然，这不是其眼睛看到的东西，而是其脑海中浮现的东西。

显然，图像分析处理领域也存在同样的问题，因此应引起关注。我们看得到期望看见的（固化在脑海中的）事物，而看不到不熟悉的事物，因此会造成误判，进行图像分析的科学工作者必须避免出现这种情况。

2.12 图像压缩

在数码静态相机和摄影机中，我们会发现关于人类视觉的一些重要信息。由于内存成本高昂，且较小的文件存储起来更加快速（远快于对执行压缩所需时间的弥补），因此所有的消费级相机和许多高端相机都是以压缩格式来存储照片的。人们似乎更愿意购买有着几百万像素的高分辨率相机，然后试图以 10、20 或更大的压缩比去压缩图像，进而生成在因特网上高效传输的文件。

针对视频的 MPEG 和针对静态图片的 JPEG 等压缩技术得到了广泛使用，且人们很少对此提出质疑。除 MPEG（Moving Pictures Expert Group，动态图像专家组）格式外，存在针对苹果公司的 QuickTime 和 Adobe 公司的 Flash 软件的多种编/解码器（压缩/解压缩器）。最初使用离散余弦变换的 JPEG 技术（Joint Photographers Expert Group，联合图像专家组）已经加入了小波和分形方法。

如第 1 章和第 3 章中所讨论的那样，所有这些方法都是通过省去原始图像中的一些信息而实现压缩的。技术上，它们是“有损”压缩技术。压缩的意图是保留足够的信息使人们识别熟悉的物体。大部分技术在一定程度上取决于人的视觉特征，并决定什么应该保留而什么可以被修改或省去。有些技术，如分形压缩，通过“借用”图像中其他位置的细节来取代实际的细节，可达到蒙蔽人眼的效果。

压缩丢弃了图像中人们不易看到的东西。人类视觉对亮度的局部突变比较敏感，而这种局部突变对应于边缘。尽管在位置和幅度上可能有轻微的偏移，但这些边缘得到了保留。另一方面，视觉上不会感知到绝对亮度，因此不会保留这些信息。由于小于百分之几的亮度变化在实际中是看不到的，甚至图像中一定距离内逐渐发生的大变化也是看不到的，而压缩技术可以在对视觉解译影响很小的情况下消除这些细节。

因为边界主要定义为亮度的变化，因此颜色信息会减少。多数压缩方案的第一步是降低颜色信息的数量，方法是对几个相邻的像素求平均，或减少图像中所用的颜色数量，或两者兼用。此外，对所有的可见光谱范围，我们对颜色的感知是不一样的。与其他颜色一样，我们不能分辨绿色范围中的较小变化。同样，颜色中的渐变如同亮度中的渐变一样，通常也是不可见的，可见的只是突变。因此，图像中颜色值的减少可能非常明显，但我们通常并不会注意到这一点。

找到图像中不变或变化不快或变化不大的区域，就可以减少视频或电影文件的大小。在某些情况下，移动物体后面的背景可被简化甚至模糊，与此同时，因为我们并不期望看到移动目标的细节，前景特征也可以被压缩。图像中引人注目（通常称为“兴趣点”，而且常与高局部对比度或熟悉的主题相关；当你看到某电影明星的照片时，你的眼睛会逗留在哪里？图像中的哪部分你不会注意到？）位置的预测，允许其他区域被进一步压缩。

当然，也许有人会说：这类压缩多数时间工作于视觉分辨阈值之下，且不会妨碍我们识别熟悉的物体。但这就是问题所在：压缩能够实现，是因为保留了足够的信息来将标签应用到图像中的物体和特征中，而这些标签反过来会导致我们使用记忆来替代图片中并不存在的细节。在取证和科学应用中记录图像的原因，不是要保留对熟悉物体和场景的回忆，而是要记录不常见的东西。如果不能事先知道哪些细节可能被证明是重要的，那么丢弃它们并不明智。若希望得到某些特征测量值（如大小、形状、位置、颜色信息等），则必须避免使用会改变这些值的有损压缩。

本节的重点并不是说在科学和取证应用中对图像进行压缩非常不明智，而是说仅在快照中对图像

进行压缩可以接受，因为人类视觉并不过于注意或依赖于一幅图像中的实际内容。识别只需很少的线索，大部分细节则被忽略。

2.13 光的世界

眼睛仅是对视觉皮质产生感官输入的整个系统的一部分，它很容易忽视光源及其颜色、位置和亮度的重要性，因此对其探讨是必要的。在购买衣服、家具或其他物品时，最好不要看它们在人工照明下的外观，而要看它们在日光下显现的颜色。光源（日光、白炽灯、荧光灯）色温的差异会使得物体的视觉颜色产生很大的不同。更不要试图在钠灯下猜测颜色，因为它基本上是单色的，不会提供任何关于颜色的线索。

在实验室环境中，如使用翻拍台或光学显微镜时，对颜色的判断同样也会受到灯泡色温变化的影响，而色温变化敏感地受施加于灯泡上的电压的影响（且随着灯丝及其表面经历物理变化后，也容易在使用的最初和最终几个小时里发生显著的变化）。调低电压来简单地减少照明（如为了拍照），可改变图像中的颜色。与钨丝光照明下所拍照片的颜色相比，荧光灯照明下所拍照片的颜色是不同的。不管图像表明光是透射的（既不反射也不吸收，如光学显微镜的透射），还是从表面反射的（如肉眼成像），都会发生这些效果。但一般情况下，大脑做好准备去解译的是后一种情况。

现实场景中可能只有单个光源（如阳光或单个灯泡），也可能有多个光源。光源可能是非常局部的，也可能是可被延伸的。光照可能是直接的，也可能是间接的，这意味着光在离开光源和到达物体的时间段内，可能会从其他表面发生反射或散射。所有这些可变因素都会影响物体在图像中出现的方式。

当然，物体的表面对其也有影响。大部分不能透过物体也不能被物体吸收的光，会从物体表面处非常薄的一层散射。对于纯金属，散射刚好发生在其表面。但大部分材料允许光透过表面下至少一小段距离。正是在这一薄层中对不同的光波长的吸收的变化，以及随着波长透过的变化，给出了一种物体的颜色。例如，对绿光优先吸收会使得物体呈现紫色。没有光透射的理想金属，不会有颜色（金、铜和银的颜色实际上源于允许微小透射的复杂电子结构）。

入射光中被反射或散射的部分可以通过表面反照率来测量。颜色非常深的物体可能会吸收多达90%的入射光，而非常明亮的物体可能仅会吸收百分之几的入射光。光和物体的相互作用通常是漫反射和镜面反射。散射成分向各个方向发出光，大致遵循一种余弦模式，如第 14 章中所述。高光成分按镜面反射特定的方向发射光，其与局部表面法线的角度等于入射角。表面的单向反射性由以镜面角反向的光与被反射光束的狭窄度之比定义。

由测量和形状信息生成表面渲染图像的计算机程序，使用对应于特定材料的模型。如第 15 章中的例子所示，改变单向反射性，是有可能改变表面的外观及我们对其组成与本质的判断的。由一组具有不同光源位置的图片，我们有可能由目标的外观来解释其几何结构。人们会无意识地这么做，因为我们的大脑会因客观世界所提供的机会不断进化，以便了解倾斜物体的外观效果及不同材料涂层的效果。

不同材料外观的变化，或光源颜色或位置的变化，有助于我们发现物体的重要细节。一部分原因是来自特定特征的反射的增强，另一部分原因是违反了我们观看表面并强制注意图像中所有细节的本意。由 Tom Malzbender 在惠普实验室开发的一种表面成像技术，在不同（已知）方向的光照下，使用由一台固定的相机拍摄的照片，来计算一个物体上每个位置的表面的方向和反照率（Malzbender et al., 2001）。图 2.47 给出了这一结果的一个例子。在增强的图像上移动光源（这通常可以交互地进行）揭示了其他方法无法观察到的细节。在这个例子中，所包含的细节是 4000 年前刻写石碑人员在湿黏土上残留的指纹，它在接近左上角的位置被识别为一组脊线。

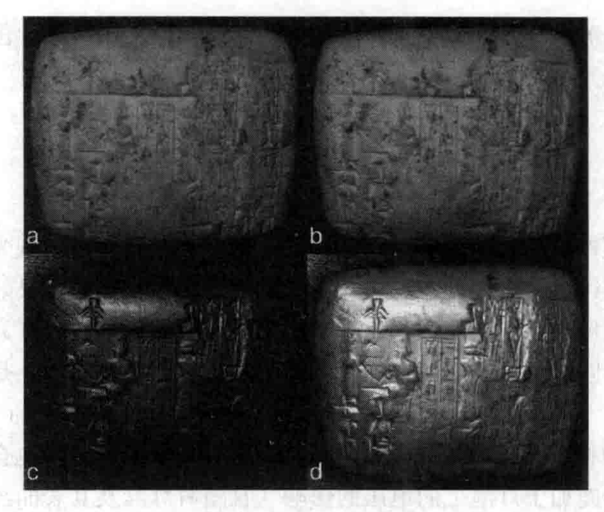


图 2.47 镜面增强：(a)有 4000 年历史的苏美尔碑的原图；(b)从计算的纹理图进行的重建；(c)提取表面法线并对每一个像素应用镜面照明模型生成的图像；(d)图(c)计算出的高亮加到图(b)上

该技术最初实现时，待拍摄物体上方置有一个圆顶并安装有每次只使用一个的许多固定光源。之后的改进（Mudge & Malzbender，2006）是在拍照现场放置一个如台球那样的黑色光亮球体，然后用便携式闪光灯从多个方向照亮该图像，这些方向由软件测定基于从球体反射回来的光确定。然后用这个数据集生成物体表面的多项式纹理图，并用任何特征去渲染出表面的图像，包括理想金属的那些特征，如同在光来自任意位置时所表现的那样。

这一计算的基本技术称为“由阴影恢复形状”或“立体视觉”。体视法从不同角度拍摄物体的两张或多张图像，而立体视觉使用来自同一视点但具有不同光照的多幅图像。由阴影恢复形状使用一种特定类型表面的漫射和镜面反射的已知分布，来估计表面关于光源和视点连线的局部斜率变化。由阴影恢复形状的缺点是，它只处理亮度（即斜率）的不同。求每个点的实际表面高程，需要以一个未知的积分常数对这些斜率值进行积分。但该方法有许多应用，也用来说明由计算机来实现大脑通过多年的实际观察被训练成自动完成的计算。

促使我们将亮度变化解释为形状的心理捷径是方便的，且通常是正确的（或至少在识别和测距等目的方面是正确的）。但它们很容易被欺骗。表面亮度可能因为几何之外的其他原因引起变化，如在物体表面上有意或无意的涂层（如氧化、污渍）。也可能存在光照不均匀的情况，如表面上的阴影。如果这些阴影未被识别和补偿，它们将影响我们对表面几何的判断。



图 2.48 对同一图像（黏土碑中的楔形文字缺口）旋转 180° 使凹点（在左边）看上去成为凸点（在右边）

我们的实际观看习惯是，期望光照来自上方，无论是天空中的太阳还是桌面上的光。若违反了这一习惯，则我们内在的由阴影恢复形状的计算将得到错误的结论并解译为凸和凹，反之亦然（见图 2.48）。当被识别后，这样的错觉是可笑的，但有时仍然会被欺骗。

因为表面的亮度随着斜率变化，就像来自阴影情形的真实形状那样，诸如扫描电子显微镜等新颖方式产生的图像看上去是熟悉的，并且很容易判断。但这涉及不同的物理过程，亮度和几何之间的数学关系不完全一样，可能产生误解。一方面，在扫描电子显微镜中，边缘和细微的突

起的外观很明亮，这不会发生在正常的表面光散射中。

许多其他类型的图像，如由 AFM（原子力显微镜）产生的基于各种针尖与样本相互作用的表面图，通常以渲染的表面表示呈现给观看者。这些图像通常是使用一个观测信号的强度作为实际表面几何来生成的，但也可能并非如此。电子或化学反应变得“可见”，就好像它们是表面的物理升高或下降。这有助于效应的“可视化”，可以充分利用我们的能力来解释表面图像，但重要（有时很难）的是要记住，这不是真正的几何，而是一些更抽象的其他属性（例如，见第 1 章中的图 1.59）。

2.14 尺寸很重要

物体的大小由边界的位置确定。对于这一相当明显的陈述来说，唯一的问题是确定边界在哪里。人类视觉基于亮度或颜色的突变位置进行定位，在图像中寻找许多类型的线来起作用，这些线包括物体的边缘。这些线被视为草图（称为“原始草图”）。漫画的原理是用所绘线条直接代替了可从实际场景中提取出来的边缘。

使用计算机程序来提取边缘线将在第 5 章中介绍。在一个常用的算法中，计算机程序发现（到最近的一个像素的）亮度发生最大变化的位置。由视网膜提取的草图却与此不同。一方面，亮度渐变不像突变那样容易看到，必须至少有百分之几的变化才能被注意到。色彩变化不能够精确定位，有些颜色的变化要比其他颜色的变化更易被注意到。

此外，人们会以不同的方式来解释边缘线。一个简单的例子是我们裁剪图案的方式，研究表明裁剪图案的方式还与性别有关。裁剪衣服的女孩倾向于裁剪在线外（接缝太宽总比太窄好）；裁剪模型飞机零件的男孩倾向于裁切在线内（使零件组装在一起）。跟踪特征的计算机测量中，也存在相同的习惯。这一差异或许微不足道，但它提出了一个有趣的问题：边缘是特征的一部分还是其周围的一部分？在许多情况下，这取决于特征在明亮背景上是亮（此时边缘很可能会被视为特征的一部分）还是暗。

在许多真实的图像中，特征的边界并不均匀。亮度和对比度的变化会导致对特征边缘的位置判断的变化，进而导致对其大小和形状的判断的变化。在某些情况下，边界在某些地方会消失（见图 2.49）。人类视觉不会被这样的裂隙干扰（但计算机测量肯定会被干扰）。我们会用简单、光滑的曲线去填补这些裂隙，这些曲线可能会也可能不会对应于特征的实际形状。

边界当然重要，但证据显明，特征在概念上并不表示为边界的集合，而表示为一根简单的中线。图 2.50 所示的毛条动物可以被识别，因为我们由所示的“骨架”填补了身体。当然，这不是实际的骨架，而是一个与计算机图像分析中所用骨架相类似的骨架。一组中线有时也被称为物体的中轴。骨架的拓扑结构（分支、端点、回路的数量）提供了由人类和机器进行特征识别的关键信息。

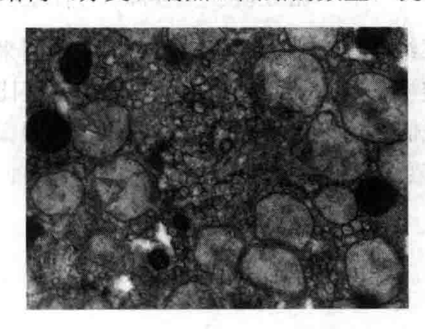


图 2.49 染色组织的 TEM 图。细胞器的膜边界在有些地方模糊不清（如箭头所示），但人的视觉“知道”用简单光滑的曲线去接续并填补它们

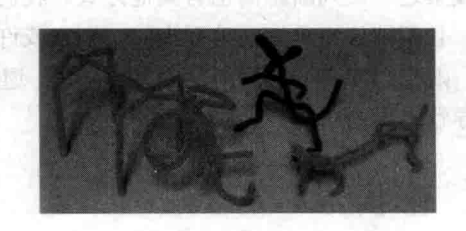


图 2.50 毛条动物（大象、袋鼠和达克斯猎狗）依据其骨架表示实体

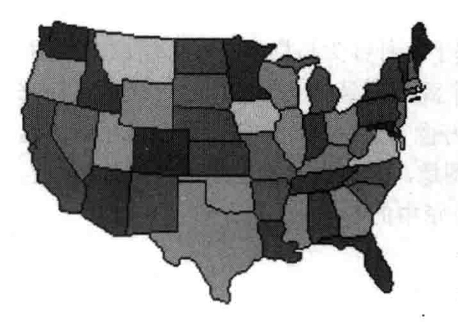


图 2.51　美国本土

不论为表示使用边界或骨架，特征的大小都会受到其形状、位置和亮度的强烈影响。美国本土地图很好地说明了这一问题（见图 2.51）。为了比较两个州的大小，我们会在脑海中准确地将一个州的图形拖到另一个的上面。由于形状不同，匹配并不完美。如何处理"突出"的部分，取决于其能感知到的重要性。例如，比较俄克拉何马州和密苏里州的大小时是很棘手的，因为俄克拉何马州柄状的狭长区域极窄而且很容易被忽视（俄克拉何马州比密苏里州稍大）。

佛罗里达州与威斯康星州差不多大（比后者小 0.4%），但它们有不同的颜色且相距甚远，所以该比较非常困难。科罗拉多州是一个简单的矩形形状，因此很难与内华达州和俄勒冈州进行比较，因为后两者的形状不是那么规则，所以看上去显得较小（实际上内华达州比科罗拉多州大，而俄勒冈州比科罗拉多州小）。北达科他州更大的垂直度在视觉上更重要，这会导致错误的结论，即北达科他州比南达科他州大。在将伊利诺伊州同爱荷华州及纽约州（后两者都比伊利诺伊州小）进行比较时，垂直度与形状和颜色上的差异也都很重要（而事实上纽约州离伊利诺伊州很远）。

即使是在最好的情形下，对物体大小的视觉判断也非常容易出错，而且很容易受到看似微小的因素的影响，其中的一些已在这些例子中得到了说明。另一种非常常见的对物体大小的错判涉及月亮。大多数人说月亮在靠近地平线时比在高空时大 1/3～1/2（见图 2.52），这可能是因为在地平线附近时存在可以用于眼睛比较的其他建筑物。垂直度通常被认为比水平度更重要（见图 2.53）。与周围环境对比更强的特征通常会被认为比具有更少对比度的特征更大。在判断大小时，几何上简单形状的偏差往往会被忽略。

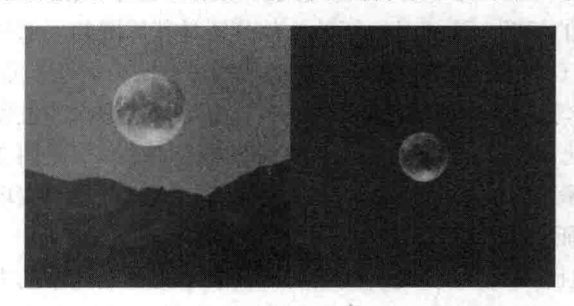


图 2.52　在接近地平线观看与在头顶上观看时，月亮的视觉增大的图解

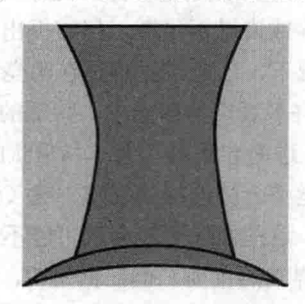


图 2.53　"大礼帽"错觉：在这顶大礼帽的夸张画中，帽顶的高度看上去要远远大于帽檐的宽度，但实际上它们完全是相等的

此外，场景的上下文也非常重要。体视法的讨论和三维解释指出，期望的恒定大小是用来判断距离的线索之一。判断距离也有其他方式。我们期望应用透视法则，因此，若场景中一个特征比另一个更高，且都被认为是搁在地面上的，则它或许因为距离我们更远而看上去应更小。如果在图像中大小相同，则我们倾向于将其判定为更大一些。遗憾的是，在观看未使用透视法则的图像时，这种"修正"可能导致错误的结论。

2.15　形状

在物体的视觉识别中，形状至关重要。即使某样东西的大小或颜色彻底改变了（如一头微小的粉红色的大象），形状仍然为识别提供了重要的信息。但形状是什么呢？无论是在英语还是在其他语

言中，几乎没有共同的形容词来实际上描述形状。有大量的词用于大小和颜色，但几乎没有用于形状的。相反，我们通过说某物的形状“像骆驼一样”来描述形状——换言之，我们并不描述形状本身，而是提供一个具有代表性物体的标签，希望听众在脑海中呈现出与我们相同的图像或模型，并且和我们一样识别出同样的重要形状特征。但有极少的明显例外，例如“圆的”这样的形容词实际上落入了相同的类别。圆表示“像圆圈一样”，而所有人都知道圆圈看上去是怎样的。

但是偏离了圆形的圈又意味着什么呢？遗憾的是，有很多种方法使圆变得不那么像圆。图 2.54 中仅给出了其中的两种：一种在水平方向伸展，但仍然是平滑的；另一种虽然保持等轴变化，但具有波浪形的边缘。其他的变化包括不平坦边缘、在更多方向上伸展等。这些特征中的哪个应该被认为“更圆”呢？

当计算机测量图像中的特征时，有很多数学方法可以用于描述形状。最常用的是第 11 章中讨论的尺寸测量的简单无量纲比值。与此相关的一个问题是所用的名字是任意的，并没有熟悉的（或一致的）的意义。这些无量纲比的另一个问题是它们不是唯一的。有可能构造很多其形状对人类来说看上去非常不同但产生相同数值测量的大量物体。但这样的形状因素对于计算机识别目的还是非常有用的。

似乎除了形状的数值特性外，人们主要依靠两类信息来形状识别：边界的粗糙度和拓扑。后者的实现方法是使用一种计算机运算来将形状简化到其骨架。第 8 章所示的骨架是特征的中线，得到它的方法通常是从边界反复移走像素，直到仅剩这样一个像素。当然，再移走像素，就会将特征分成碎片。轮廓（边缘）和骨架是第 12 章中所示的一些经典程序的简化目标描述。

即使是对特征骨架或边界的不完全划定，通常也足以构造一个视觉印象（可能是正确的，也可能是错误的）。这种能力可仅使用几个关键特征（角、端点和分支点）来表征几种常见错觉的特征形状。Kanisza 三角形（见图 2.55）是在我们的头脑中构建的，方法是将三个定义很好的角点连在一起。这种连接总是平滑的，但不必是直线。

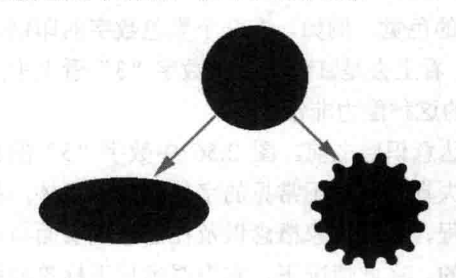


图 2.54　下方的两个图形与上方的圆面积相同，但形状不是圆的

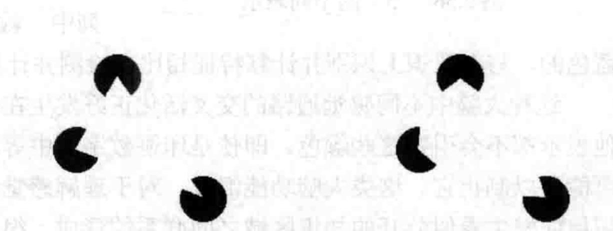


图 2.55　Kanisza 三角形是用直线或柔和的曲线将角标记点连接起来形成的一个幻觉区域。由此形成的三角形内部可能看起来比周围环境明亮

一旦建立起了对象形状的基本拓扑形式，人们直觉上识别形状的第二个性质就是测定边界的平滑度或不规则性。因为自然界中很多东西具有分形几何而不是欧氏几何，这表现为在边界分形维数中检测差别的能力。分形维数的测量将在第 8 章和第 15 章中讨论。记住，当人类可以用不同的边界不规则性比较性地对物体分级时，他们并不在视觉上进行测量。

传达识别物体形状特征信息的广为接受的方法，仍然是向他人显示物体的一张图片。对于非常相似或仅有的差别在正常观看尺度下不可见的近似特征，这种方法效果很好。遗憾的是，在大多数科学中，感兴趣的物体，无论是材质上的缺陷、幻灯片上的癌细胞，还是虫子的新物种，并不完全一样。尽管用来识别的线索（若被正确地选择）仍然存在（尽管并不一定在每幅图像上都可见），但自然变化是明显的。

在展示“有代表性的图像”时，科学家们会尝试向同事传递这些线索。图像几乎总是需要大量的

文字作为补充。但是，如果同事不熟悉这些对象（这当然是需要沟通的原因），他们能选出同样的特征或线索吗？而且选出来的图像足以表示自然物体的变化范围吗？或者足以表示那些变化之间的相关性吗？这些危险经常出现在使用“典型图片”时，即使图片确实给出了正确的表示；此外，在大多数情况下，可能需要承认图像被选中不是在分析后证明它在任何统计意义上具有代表性，而是因为该图像满足了一些未表达出的其他审美标准。单幅图像仅能提供的轶事证据，说得乐观点是冒险，说得严重点会有误导性，因此只要可能就需给出定量数据。这是使用计算机来成像的另一个动因：获得足够多的测量数据来进行有意义的描述。

2.16 上下文

物体的识别通常受其上下文的影响。有时上下文由图像本身提供，但其更常来自先验知识或独立信息。图 2.56 中有很多非常不同的数字“5”的表示，包括用我们不懂的语言书写的“5”。但是一旦“5”这个概念被接纳，这些不同的表示就都变得可以理解了。这一过程涉及了与图像无关的很多知识，它发生在比基本形状识别更高的思维层次。知道五角形有五条边可能有助于我们翻译希腊文，回想五月五日节（Cinco de Mayo）[8]派对可能有助于西班牙语等。



图 2.56 “5”的不同表示

对该识别过程的一个有趣洞察，来自于对受到“共感”折磨的患者的研究。“共感”是一种现象，即大脑中的某类交错会将来自某种感知的输出与另一种相混淆。例如，具有共感的人可能报告在钢琴上弹奏的某个特定音调可能引发特定的味觉。在最常见的共感形式之一中，看一个数字会引起特定的色觉。例如，在一个黑色数字的印刷阵列中，数字“5”看上去是红色的，而数字“3”看上去是蓝色的，与在意识上识别并计算特征相比，检测并计算特征的这种能力非常快。

这种大脑中不同感觉通路的交叉活化正好发生在信息抵达意识层之前。图 2.56 中数字“5”的其他表示都不会引发这些颜色。即使是印刷数字的中等程度的失真，例如不常见的字体或轮廓字体，都可能足以制止它。这类大脑功能混乱，对于理解感觉信息过程、提取抽象概念以及在潜意识层面与意识层面发生看似分开的知识区域之间联系的形成，很是重要。在此情况下，它表明远早于标签前就发生了基本形状识别，且特征上应用了语义内容。

在同一张图片里甚至可能有多个上下文，特别是在阅读单词时。由于通常有足够的冗余和上下文来理解信息，因此我们能容忍拼错及潦草的笔迹，即使图片本身是错误的、不完整的或模糊的，如图 2.57 中的例子所示。

It's hard work digging clay
Save it for a rainy day.

图 2.57 每行最后的几个单词的形状相同，但我们可以通过由其他单词构成的上下文正确地阅读

上下文对于校正科学图像中的数据的解释至关重要。我们基于先前的经验及学习、关于样本如何准备以及图形如何获取的知识的期望，甚至对于一个经验如何形成的希望或恐惧，都会明显地影响特征的视觉

⑧ 五月五日节是墨西哥的一个地区性节日，主要为纪念 1862 年 5 月 5 日，墨西哥军队击败法国殖民军，取得普埃布拉战役的胜利，因此也是墨西哥传统的爱国主义节日。——译者注

解释和识别（或不能识别它们）。为了能正确地解释它，要对样本有足够的了解，同时避免期望所引起的陷阱。有些人擅长于此，而且任何人都可能会偶尔出错，但幸运的是，科学出版物的公开性提供了纠错的机制。

上下文的一个常见问题出现在显微镜学中。通常横贯组织的截片被放在显微镜下进行透射检查，或抛光不透明材料的表面来进行反射检查（或等效地使用透射式电子显微镜或扫描电子显微镜）。在所有这些情况下，图片本身是二维的，但是它们所表示的结构及样本是三维的。大多数人很难提供一个合适的上下文来以三维结构理解这些图像。第 9 章中所涉及的体视法提供了可以用来将二维图像上的适当测量联系到三维结构的几何关系，但很少有观看者能以足够的直觉在视觉上正确地解释这些图像。

即使是在强化训练且熟悉特定的三维结构后，这看起来仍然正确。医生依赖来自诸如核磁共振成像（MRI）和 X 射线计算机断层成像（CT）这样的设备的剖面图来分析人体。但是由测试和采访显示，几乎没有人在脑海中具有该结构的三维图像。相反，他们学会了识别几乎总是取几个标准方向的截面的标准外观，并学会了识别与基准的偏差，尤其是与常见疾病或其他问题相关联的偏差。

2.17 必须做好布局

人类非常擅长的一件事情就是寻找物体布局的规律。有时，这种对简化的探索会发现目标之间真正而有意义的关系，但有时却并非如此。纵观历史，很多文明都能通过连接天空中的亮星来定义星座（尽管结果不同）。图 2.58 显示了古典希腊的版本。搜集需要用于构造巨石阵的数据作为夏至、冬至、日食、月食的预测，必然会耗费数代人的时间。允许椭圆来简化越来越复杂的托勒密行星运动的圆环-外圆环模型的哥白尼革命，是对这类简化的一个探索。

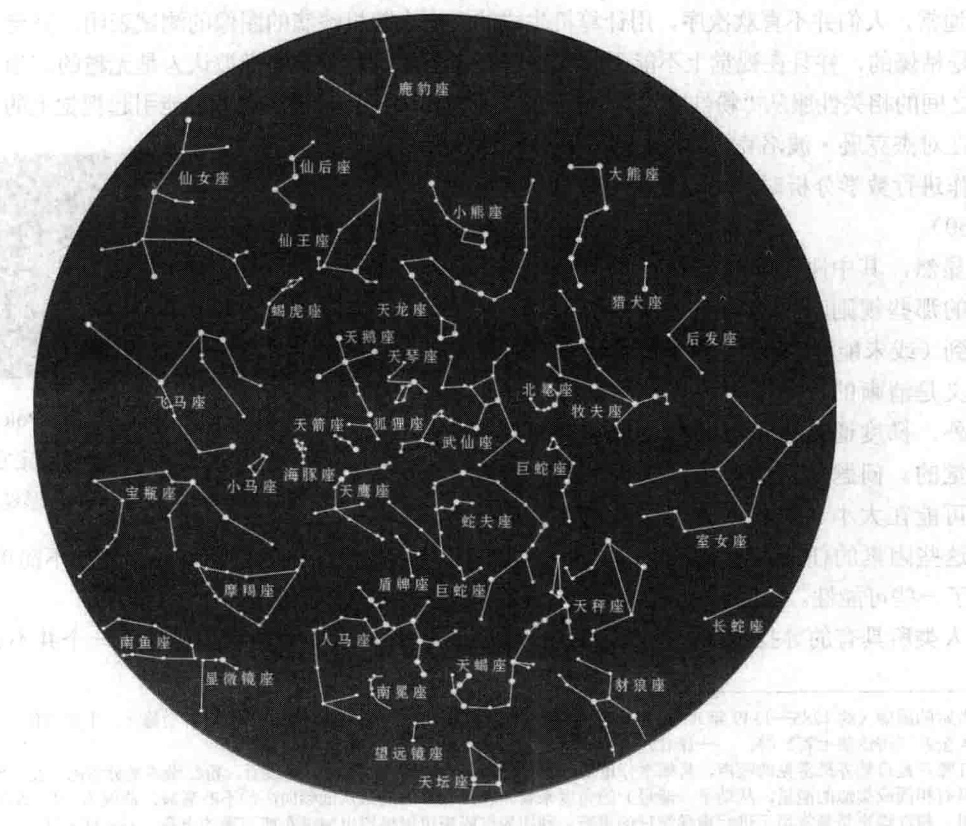


图 2.58 熟悉的恒星星座

大多数科学家都会遵循爱因斯坦的格言，即重要的是找到有效解决方案中最简的一个，而不是太简单的那一个。这种想法早就有之。奥卡姆的威廉[9]的“吝啬原则”就是“如无必要，勿增实体。”凭直觉我们都能寻求到简单的答案，但有时却一个也找不到。当然，这也适用于图像检查。

寻找图像或数据图中视觉对齐的点时，人们更倾向于直线或平滑、渐进的曲线。线性回归可能是广泛用于对一组点强加规律（并滥用）的数据解释方法。在线条或边界上填充裂隙往往是一个有用的过程，但也可能会导致错误。图 2.55 中产生错觉的 Kanisza 三角形是一个填补裂隙并连接点的例子。连接点和线的过程与前面所讨论的分组密切相关，如在色盲测试图像中从色环形成一个数字的过程（见图 2.30）。

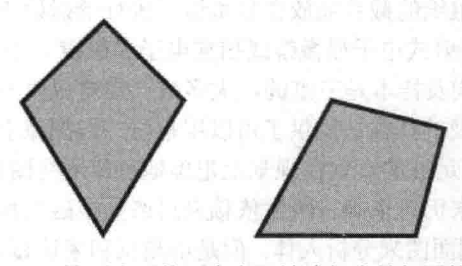


图 2.59　这两个形状不会立即被感知为是相同的。左侧的“风筝”有一个明显的垂直对称轴。右侧的不规则四边形有水平的底部

人类视觉存在固有的方向性偏好，即最喜欢垂直方向，接下来是水平方向，且钟爱于对称性。如图 2.59 所示，它会使我们出现误判。它还会影响我们检测梯度或聚类的能力。

有些处理会产生特征的随机分布，如同在桌面上溅洒盐一样。若每个特征完全与其他特征无关，就会得到一种随机分布。但由于特征相互之间要么吸引要么排斥，因此会出现非随机的分布。生长在沙漠中的仙人掌是自回避的，因为每一株仙人掌都试图保护其水分和养料的供给。由于表面张力效应，漂浮在液体表面上的微粒可能聚集。学生在教室里是自回避的（等间隔的），但在聚会时是聚类的。在极端情况下，目测聚类或自回避是有可能的。多数情况下，人们不容易识破表面的混沌分布以便检测其效果，但对数字图像的测量（在第 10 章中涉及）可以经常提供定量的描述。

通常，人们并不喜欢次序。用计算机生成的表面上随机漆滴的图像的测试表明，完全有序的图像被认为是枯燥的，并且在视觉上不能引起兴奋，而完全随机的图像同样被认为是无趣的。当大小、颜色和位置之间的相关性服从“粉红噪声”[10]或分形关系时，该图片对于观看者最能引起视觉上的兴趣。有趣的是，在对杰克逊·波洛克（Jackson Pollock）[11]的几幅画作进行数学分析时，可以发现同样的关系（见图 2.60）。



图 2.60　杰克逊·波洛克的“Blue Poles #11”。已对其画作进行了分析，以确定它们具有其生涯中逐步形成的一种分形结构及复杂性

显然，其中刚好有我们在视觉上必须努力去找到的那些规则暗示的分布是最吸引人的。我们注意到（或未能注意到），科学图像中的特征布局的含义是清晰的。除了可能在整个图像中呈现的布局外，梯度通常也是重要的，但它并不总是容易察觉的。问题是可能有很多不同种类的梯度。特征可能在大小、形状、方向、颜色、密度、数量或这些因素的任意组合上，作为位置的函数变化（或梯度变化的空间路径可能并不简单）。第 9 章中说明了一些可能性。

人类所具有的寻找图像中的次序的先天能力是有风险的。有时我们想象的一个并不存在的规则性

⑨ 奥卡姆的威廉（约 1285—1349 年），出生于英格兰萨里郡奥卡姆，14 世纪逻辑学家、圣方济各会修士。主要著作有《箴言书注》、《逻辑大全》、《辩论集七篇》等。——译者注

⑩ 粉红噪声是自然界最常见的噪声，其频率分量功率主要分布在中低频段。从波形角度看，粉红噪声是分形的，在一定的范围内音频数据具有相同或类似的能量。从功率（能量）的角度来看，粉红噪声的能量从低频向高频不断衰减，曲线为 $1/f$，通常为每 8 度下降 3 分贝。粉红噪声是最常用于进行声学测试的声音。利用粉红噪声可以模拟出如瀑布或下雨的声音。——译者注

⑪ 杰克逊·波洛克（1912—1956 年），美国画家及抽象表现主义运动的主要力量，以其独创的滴画而著名。——译者注

或次序（如星座）、复杂性或表面的无序存在，实际上隐藏了真实的深层结构。某些方向可能比其他方向更容易检测，而复杂的梯度有可能逃过检测，除非我们事先知道所要寻找的是什么。因此，许多真实的空间布局可能会错过，即使在二维空间中也是如此（因为通过检查穿过三维结构来检测二维截面引入了其他的问题，这个问题对于三维空间布局更加糟糕）。

2.18 眼见为实

若人类视觉在对图像的解释上如此轻易地被欺骗，则影响人类对于结果判断的不同图像处理程序的效果的问题就有可能（并且是经常地）会出现。

这在某种程度上是一个伦理学问题，所处理的可能是有意构造来误导观看者的合成品或赝品。该问题与任何其他故意欺骗的情况是完全相同的。如果有人愿意耗费足够的时间和技能创造一个赝品，是可以做到的，无论是作为图像或其他类型的数据，以及是否使用了计算机（许多优秀的伪造图像已在摄影暗室里精心制作，更不必提赝品画作了）。

检测这样的赝品可能非常困难，在某些情况下也许并不可能。在特定的数字图像情况下，用于剪切和粘贴的软件工具可使任务更为容易，但结果通常很容易会被辨认出。检测出图像中的变化或一幅图像到另一幅图像的粘贴部分存在大量的线索。一些明显的情况是阴影长度和角度不匹配、几何形状不一致、焦点变化、颜色不匹配以及不同的纹理或信道噪声（包括噪声随亮度变化的方式）。



图 2.61　对两幅不同的图像的部分进行组合的例子

在图 2.61 的例子中，一只德国牧羊犬的头部被粘贴到了一只爱尔兰塞特犬的身体上。在这种情况下，很容易检测伪造，但注意到所呈现的各种线索是有用的。第一条线索当然是狗脖子上颜色的相对突变。可能已经进行了一些颜色的调整来减少差异，边界有时可能是隐藏的（例如在上面放置了一个颈圈）。在这种情况下，执行一些模糊处理来防止清晰的线引起人们的注意，但局部模糊的存在及精致细节的减少本身就是一幅图像已被做了一些手脚的线索。

该合成图像的第二个常见问题是，图像中两部分的光照是不同的。狗的身体是从图片的右侧照明的，而其头部是从前面照明的。同样值得注意的是，狗的影子与合成物不一致，它仍然具有原头部的形状。

头部图像原本较小，只能放大并旋转来与身体对齐。这导致了与原图不匹配的像素化。精细细节的比例在图像的两部分中也是不同的，这表明它们并不是一起的。通常，用不同的相机在不同的光照条件下拍摄的图像有不同的噪声特性。在红色、绿色和蓝色通道中的噪声比及噪声振幅随着绝对亮度的变化，需要更多的努力来测量，但在对来自不同来源的图像进行组合时，几乎不可能进行匹配。

用足够的努力（和技巧），可以实现无缝的和事实上无法察觉的结果。然而，在科学研究中也有少数情况，即其中一幅单独的图像对于结果提供了确定性的证据。创建记录研究学习的一组虚假图像很可能是一项不切实际的艰巨任务，而在任何情况下科学方法的特点是能被他人重复的实验能力，所以故意伪造的图像通常并不是严重隐患。这种问题更经常地出现在取证的情况下。

为了科学和技术地工作，最好能将原图与那些为印刷、观看或测量而处理的复制品区分开，并且保持处理步骤的详细记录，以便能在出版物中报道及需要时重现。可以配置一些软件包（如 Adobe Photoshop）来自动地完成这些工作，并在磁盘上保存图像所有处理操作的历史。在大多数的取证工作中，这被认为是必不可少的。

更令人关注的是在不知不觉中引入的偏差，这些偏差会通过选择所拍摄区域或剪裁照片以隔离感兴趣目标而产生。小报常常被挑剔使用了断章取义的名人图片，但即使是负责任的出版物，也必须决定移除背景和周围环境是否对故事合适或相关，而且这也只可能在这种关联性变得明显之后。

在科学中，故意伪造可能很少见，即使有也通常会被校正，因为科学的方法要求其他人能重做并验证结果。如果科学家发表了一幅图像而没有关于过程的足够信息，就会受到批评。因此信息通常就在那里，而具有可比技能水平和知识的其他人可以作为检查者。这并不意味着每一种出版物都应该包含这样精细级别的细节，以便使不了解应用或图像处理过程的任何人都能够照着做。

开放性及取决于自动确认与再现性的科学方法处理蓄意篡改的问题。这些通常对图像处理没有价值，而对选择图像、标本及解释结果更有用的无意识误差，实际上更受关注。当然，这些可能也涉及人类视觉系统，但通常本质上是误解实验背景的结果。

记住，本书中讨论及说明的有关图像处理技术，并不向图像中添加任何内容。相反，通过移除一些图像内容，这些技术使得在视觉上更容易接近剩下的细节，这些细节在特定的上下文中是大概相关且令人感兴趣的。因此，消除或抑制阴影、噪声或高对比度，可以揭示出之前隐藏的细节。或者，增加局部对比度，或将纹理变化转换为亮度，或用一种颜色取代另一种颜色，可以看到低于可视阈值的细节。

2.19 结论

人类视觉是一个极其强大的工具，经过数百万年的进化后，人类视觉可从场景中提取那些对于我们作为个体及作为一个物种生存来说非常重要的细节。视觉信息处理通过结合高度并行的神经回路来检测并关联图像内特定类型的细节。许多“大部分时间”有效的捷径被用于快速识别。研究这些在各种视觉错觉中显示的技巧的失败，有助于理解基本的处理。

对于依赖于图像的视觉检查来获取或解释数据的科学家而言，认识到失败和偏差同样是重要的。视觉检查只是一种比较的过程，而不是一个量化的过程，它很容易因图像中其他信息的存在而引起偏差。克服绝大多数这些特定问题的计算机图像分析方法是存在的，但它只提供与所提问题几乎一样好的答案。大多数情况下，如果科学家不能在视觉上感知出原始图像中的特征或趋势，后续的测量将无法进行。

几本经典的书籍对人类视觉提供了很好的介绍，但它们深入研究关于视觉皮层解剖细节的专业文献：

John P. Frisby (1980) *Illusion, Brain and Mind*, Oxford Univ. Press

John P. Frisby and James V. Stone (2010) *Seeing, 2nd Edition: The Computational Approach to Biological Vision*, MIT Press, Boston

David Marr (1982) *Vision*, W. H. Freeman Co.

Irvin Rock (1984) *Perception*, W. H. Freeman Co.

当然，在许多同行评议的期刊里也有大量的文献；今天，任何人都不应忽视互联网搜索，这会找到该主题的一些课程笔记、重印出版物和良莠不齐的网站。

第3章　打印和存储[①]

创建图像的硬拷贝表示，例如用做报告和出版物中的插图，对于图像处理系统的许多用户而言是很重要的。对于存储图像，这通常也很重要，因为可在以后检索它们，例如为了与新图像进行比较或为了传送给他人。图像的打印和存储是硬件及软件开发的活跃领域，经常会出现重大进展。而照相机和扫描仪的发展，主要动力并不是科学、技术和取证应用的需要。

打印和存储是必要的，因为几乎不可能将图像简化为简洁的语言描述，或对其进行一系列的测量，进而将某人所见或所信的图像中的重要内容传达给别人。事实上，基于对一幅图像或多幅图像的检查，让他人注意到图像中具有显著特点的具体细节或一般结构，往往是很难的。面对无法找到描述性文字或数字的情况，人们采用的方法是传递可能带有注释的图像。

3.1　打印

本书英文版是以彩色印刷的，采用了通常不面向个人图像处理用户的商业印刷技术。这种类型的印刷设备自身就存在一些挑战。但使用相当廉价的机器也可以很好地处理许多日常工作，因为单色和彩色打印机的质量、速度一直在不断提升，而成本一直在下降。典型彩色激光打印机的价格是几百美元，它已成为台式计算机系统的常用外设，尤其是办公室或实验室中小型网络的常用配置。打印机主要用于打印文本和简单的图形，但其中的多数打印机同样可打印出令人满意的图像。喷墨打印机（价格低廉，可在购买计算机时低价购买或免费得到，但所用耗材昂贵）也能很好地打印图像。使用染料升华或照片印刷方法的服务部门，可提供在线连接和快速转换。自从在电传打字机上使用 X 和 O 来表示不同灰度级，进而打印圣诞海报的计算机图形出现之后，技术已有了相当大的发展。本章将研究用于个人计算机图像处理系统中的图像打印技术。

因此，打印机是否使用诸如 PostScript 这样的高级页面描述语言并不重要，只要它们允许计算机传输给打印机一个单独像素亮度值的数组，PostScript 就能被用来以最大打印机分辨率产生平滑的字符和线型。大多数打印机除了可以用这种方法控制简单的文本打印外，还可以创建任何图形输出。这意味着“菊轮”或其他形式（现在已过时）的字符打印机不能用于成像。这就是电传打字机的工作原理，即使用字符套印技术来产生不同的灰度。但使用色带在纸张上以点阵形式进行输出的点阵打印机、喷墨打印机、热敏式打印机和其他设备，也可用于打印图像，打印结果的质量主要取决于点的大小和点间距。

将几种颜色的几个点放在一起，可产生视觉上混合并感知为连续色调彩色图像的思想，最初由画家用来创建称为“点画法”的绘画风格，其中著名的画家有乔治·修拉（Georges Seurat）[②]（见图 3.1）。CRT 和 LCD 显示器中使用了同一原理，它使用红色、绿色和蓝色的组合来形成整个色域。打印原理同样如此，不论是打印本书的打印机，还是大多数台式打印机。印刷使用的基本颜色是青色、品红色和黄色，它们与白纸一起可产生一系列的颜色。

① 本章的部分内容最初以一系列文章出现在 2004 年皇家显微镜学会会议论文集（*Proceedings of the Royal Microscopy Society*）中。

② 乔治·修拉（1859—1891 年），法国新印象主义的重要代表人物。——译者注

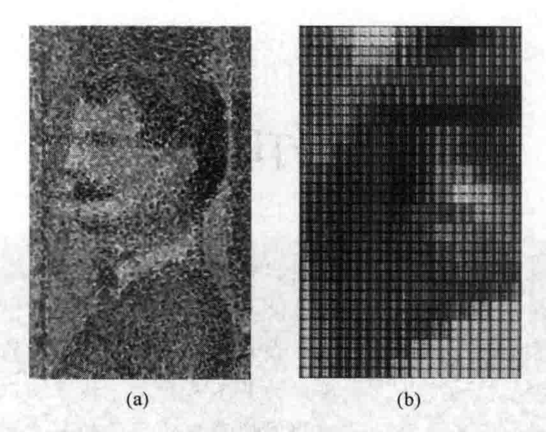


图 3.1 点画法：(a)Georges Seurat 的“La Parade”（1889 年）的细节，显示了用于创建连续色调的漆点；(b)显示各个红色、绿色和蓝色的 LED 显示器的放大图，这些颜色亮度的变化经视觉组合就产生了连续颜色的感知

在以图像质量解释打印机指标时，我们经常会混淆“点数/英寸”（dpi）和“线数/英寸”（lpi）的概念。对于上面提及的任何打印机，尤其是激光和喷墨打印机，dpi 是指打印机可淀积在纸上的小黑（或其他颜色的）点的数量。尽管有些打印机的水平和垂直 dpi 分辨率不同，但在纸张上两个方向是相同的。一般情况下，点用于形成字符和线条。dpi 很小时会使得字符看上去粗糙，且线条看起来呈阶梯状或“混叠”在一起。当前 1200 或更高 dpi 分辨率的激光打印机是很常见的，这种打印机打印出的报告文本、图形能令人接受。喷墨打印机通常以每个圆点中的液墨量来指示圆点的大小，其典型值通常为几皮升③。这些点小到几乎看不见，因而结果图像就呈现出了连续的色调。

但这些打印机放置在纸上的点由黑色或几种颜色组成。典型的彩色激光打印机只使用黑色和三种彩色墨粉（青色、品红色和黄色），喷墨打印机至多使用 7 种墨粉（如黑色、灰色、黄色、品红色、浅品红色、青色和浅青色）。打印连续范围内的颜色时，由于单个点的亮度和大小不能变化，因此有必要组合使用这些点，这种技术通常称为半色调。这种技术常用于报纸、杂志和书籍，也可用于彩色和单色图像。报纸和书籍间半色调质量的区别，主要在于纸上每英寸能放置的点的数量，以及生成灰度或彩色结果的组织方式。

灰度（单色）半色调的基础是打印机产生的各个黑点进行组合。例如，16 个点组成的 4×4 阵列可形成表示一个像素的单元。在该单元中，部分或所有的点被打印出来。在未打印出点的位置，仍为纸的颜色。如果单元很小，那么观看者将看不见各个点，但能在视觉上平均暗墨和亮纸的数量，进而形成一个灰度。如图 3.2 所示，在该例子中，从纯黑色（所有点被打印）到纯白色（没有点被打印）的范围内，有 17 个可能的灰度级。

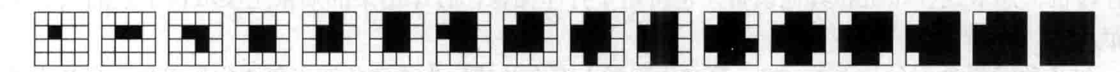


图 3.2 用一个 4×4 点单元产生的半色调灰度。在所有位置、部分位置打印点或在全部位置不打印点来产生 17 种不同的灰度值

300dpi 的打印机会在纸张上以 1/300 英寸的间隔于每个方向上放置各个黑点。将这些点组合成 4×4 半色调单元，可产生每英寸 300/4 = 75 个单元。若每个像素对应于一个半色调单元，那么图像可以与其在屏幕上显示的相同尺寸进行打印（典型的显示器分辨率范围为 72～100 像素/英寸，笔记本电脑显示器分

③ 1 皮升等于十亿分之一公升。——译者注

辨率可能更高)。每个半色调单元使用 17 个可能的灰度级之一来表示像素的亮度。图 3.3 说明了使用这些打印的单元来表示平滑文本中灰度级的方法。

当然，由于原图像通常可能有 256 级灰度，17 级灰度是相当差的亮度信息表示。但这不是唯一甚至也不是最严重的限制。宝丽来（Polaroid）胶片的立等打印只显示了大约相同数量的不同灰度级（胶片的底片要远好于打印件），但这对许多科学目的是有用的。

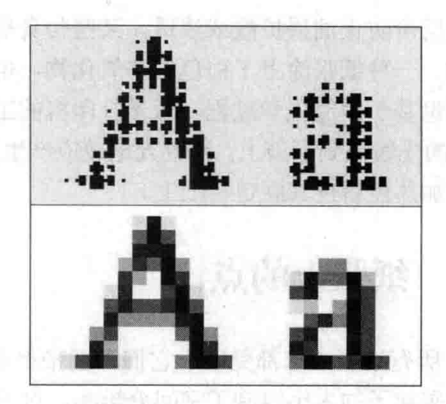


图 3.3 用半色调显示平滑字母的灰度。打印文字时，字母通常打印为纯黑色，这说明了半色调用于表示图像中的亮度值

上述半色调方法的一个问题是，点和单元大到足以被观看者在视觉上进行区分。在杂志或书籍中，单元的尺寸比较小。单元大小通常被描述为网屏或网格：以线数/英寸（lpi）表示的网屏间隔通常对应于上述的单元数量/英寸（但有时也以包含了一个分辨率元素的行对数量/英寸给出，因为它是两行，一行亮一行暗）。超过 100 行/英寸（通常为 133 行/英寸或甚至高达 175 行/英寸）的网屏被用来生产高质量的印刷插图。在报纸插图中，通常会使用至少 85 行/英寸的网屏。图 3.4 显示了一台激光打印机的半色调输出的几个例子，它通过改变半色调单元的数量来平衡灰度和横向分辨率。当前的主流台式激光打印机足以胜任报表输出，但还不能胜任出版物输出。此外，通常很难再现这种打印件，且会受诸如莫尔条纹或干涉图样这类问题的困扰。

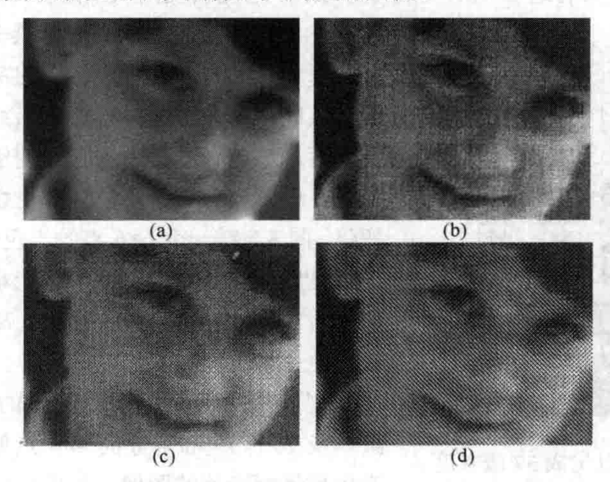


图 3.4 使用激光打印机打印出的半色调图像的放大细节：(a)原始图像；(b)100 线/英寸图像；(c)75 线/英寸图像；(d)50 线/英寸图像。增加半色调单元的数量时，以可见灰度级数为代价改善了横向分辨率，反之亦然

要保持图像中特征的清晰度，不仅需要更多线数/英寸的分辨率，而且需要由这些良好分隔的单元来表示更多的亮度级。这意味着打印机（或激光照排机）必须能够放置大量的小点。在一个单元中使用 8×8 阵列可形成 65 级灰度。使用 125 线网屏时，这对应于 8×125 = 1000dpi。这可能是用于商业图像打印的排版质量的最低值。（将在下面介绍的）色彩会施加需要更高分辨率的其他限制。高品质印刷所用的激光照排机的分辨率一般能够达到 2400～3000 点/英寸。

上述半色调方法的另一个问题源于点自身。不同的印刷方法会以不同的方式产生点。点阵打印机使用很小的针头对着纸张敲击色带。喷墨打印机产生微小的墨滴。这些打印机中，有些打印机的油墨以液体形式渗进纸张，形成稍微模糊的点，有些打印机的油墨凝固并附着在纸张表面。高级纸张的表

面涂层可防止油墨扩散或渗透。某些传真机仍在使用的热敏打印机，使用一根针来使电流穿过纸张的涂层。一种纸张涂上了白色的锌氧化物，电流会感应涂层而在相应位置上沉淀出金属的暗斑，其他热敏纸也基于银的化学过程。激光打印机的工作原理本质上和静电复印机相同。来自激光（或光电二极管）的光线落到硒鼓上，通过光电效应产生局部静电荷，静电荷拾取碳粒子（色粉），使其转移到纸上，然后加热色粉使其固定在纸上。

3.2 纸张上的点

所有这些技术都受限于它们在纸张上形成微小暗点的能力。在复印机或激光打印机墨盒中用做色粉的碳粒子的大小限制了空间分辨率，对于 1200dpi 及更高数值的分辨率，需要特别精细研磨的色粉。制造更高分辨率激光打印机的主要限制，并不在于打印机所需的额外内存，也不在于需要将光聚集在硒鼓上的一个很小的点上，而在于色粉颗粒的大小。有些系统会使用液状容器使色粉散开，来精密控制色粉微粒的放置。

类似的约束限制了其他印刷方法。使针尖敲击色带而在纸上沉淀小而暗的墨点的困难，或热敏打印机所书写出的圆点的模糊性，阻碍了这些技术向更高分辨率的发展。喷墨打印机可以产生小液滴来沉淀小圆点，但油墨极易在纸张上扩散。纸张表面的粗糙，需要特殊涂料来防止油墨浸入纸张纤维或散布到整个纸张表面，或者所放置色粉颗粒的减少，都成为关键问题。除了购买高质量的打印机外，还需要对特定打印机使用具有适当光洁度的特殊纸张，才能实现高质量的打印（Lee & Winslow，1993）。

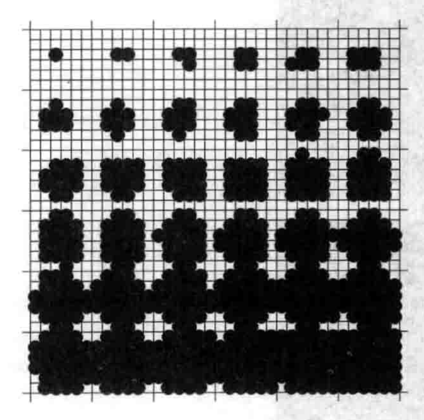


图 3.5　6×6 点的半色调可以生成 37 级灰度。使用大到足以触及对角线的近似圆点会使它们重叠并产生更为理想的暗单元

由于打印机生成的点通常具有瑕疵且边缘粗糙，因此很难保证单元内的点阵所生成的灰度是一致的。大多数点要比其间距大，因而可能打印出纯黑色区域。对于打印字符而言，这是利好，因为它们期望是纯黑色。但这意味着点会在单元内与一些白色区域重叠，进而使得半色调灰度变暗。图 3.5 对一个 6×6 点的半色调单元演示了这种情况。在尺度的暗端，相邻灰度级可能难以区分，而在亮端，最初几级之间的差异可能会很大。高质量的图像印刷通常会设法避免纯白色和纯黑色。

对于上文所述使用 300dpi 的打印机、17 个标称灰度级以及 75 个单元/英寸的 4×4 点单元的情况，灰度变暗会产生质量相当差的图像。如果打印机的分辨率高到可以创建具有更多阶的更加精细的半色调单元，就有可能构建一个（将像素灰度值转换为打印灰度值的）映射来校正图像的变暗。这些调整曲线在软件内部使用，可在页面上打印或多或少的等阶亮度。控制计算机到打印机的信息流的软件驱动，会处理很多困难的工作，包括对每种色墨或色粉应用非线性校正。必须为所用油墨和纸张的组合应用称为 ICC（International Color Consortium，国际色彩联盟）的特定校正曲线。这些曲线可能由每个系统的特定校准产生，也可能由制造商提供（在后一种情况下，使用第三方的墨水和纸张时可能会产生意外的偏差和很差的结果）。

由于人眼对灰度的响应不是线性的，而是近似对数的，因此产生视觉上亮度适当的打印图像就需要进一步调整显示曲线。由于这些限制，使用由 8×8 点半色调单元定义的 65 个灰度值的打印尺度，可能仅够显示出所打印图像中的许多明暗度的一半。

在排字机或激光照排机中的半色调网格不受点的大小或完美度的限制。排字机使用光线而非粗糙

的色粉微粒或蔓延的墨滴来对感光乳剂进行曝光，然后将其显影以制成胶片。胶片乳胶中银粒的大小要比有效的点尺寸小很多，因而可以在大小和形状上非常精确地控制点。但仍需要调整曲线来弥补人类视觉的非线性响应。

如果整幅图像的亮度变化是平缓的，且灰度级数很少，那么就有可能会导致一种称为条带或色调分离的视觉效果。图 3.6 说明了这一点。一个亮度级到下一个亮度级的界线在打印的图像中显示为轮廓线，且在视觉上它被识别为图片中的一个特征，而它其实纯粹是印刷工艺的一个缺陷。增加所用灰度级的数量，或在某些情况下修改半色调单元内产生灰度的点的排列方式，可以避免条带效果。一般情况下，最大数量的灰度级数是 $1 + (\text{dpi}/\text{lpi})^2$，其中 dpi 是单位为点数/英寸的打印机分辨率（例如，对于典型的激光打印机，其为 1200 dpi），而 lpi 是半色调网屏的线数/英寸（例如，对于上面的例子是 75 线/英寸）。这里假设可以使用所有的灰度级，以便满足上面提到的变暗和视觉效果等条件。

在上面的例子中，在白框内产生或多或少的圆形黑色区域时使用了点的一种排列。在每个单元使用较大的点阵，可构造出更规则的圆形图案。圆点图案是最常用的排列之一，但也存在许多其他的排列方式，包括线和交叉。每种排列方式都会产生一些视觉缺陷，这些缺陷对于达成某些艺术效果可能有用，但通常会对我们难看图像内容造成干扰。例如，多数 PostScript 打印机会使用菱形图案来近似圆点，这会导致密度小于 50%黑色的点，看上去完全不同于那些密度为 50%黑色或更大的点。原因是，在 50%的密度处，相邻单元的点会相触，使得眼睛感知到从白色背景上的黑点到相反情况的突变，即使各个点看起来并不明显（见图 3.7）。所有这些缺陷都会退化原始灰度图像的表示。

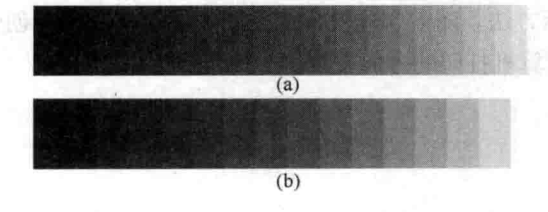


图 3.6 显示了条带的灰度条：(a)32 个灰度阶；(b)16 个灰度阶

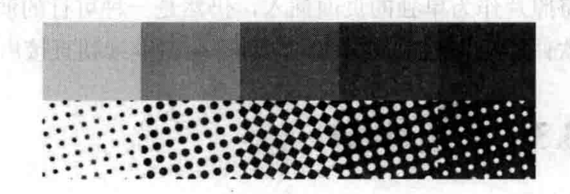
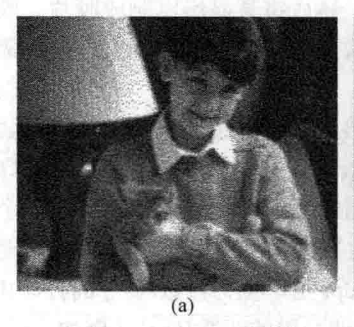


图 3.7 当覆盖率达到 50%时，半色调点相互接触，将效果从白色背景上的黑点改为黑色背景上的白点

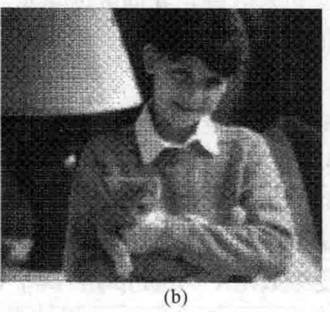
多数喷墨打印机会在单元内随机放置各个墨点，因此单元大小和对齐看起来并不明显。且当每个半色调单元中的点随机分布时，似乎会出现更少的视觉缺陷。这就需要在打印机内部进行其他处理。其他打印机使用分析邻域单元的程序。若一侧的单元暗而另一侧的单元亮，则该程序可能会使用一种向着暗邻域移动的点模式来产生更平滑的边缘。在为印刷字符产生平滑的边界及最大限度地减少线条上的锯齿边缘时，这种程序的效果很好，但不适用于灰度图像的印刷。

中等 dpi 分辨率的打印机不能产生很小但仍能表示许多灰度级的半色调单元，如果在单元内使用点的规则图案，那么在图像中可能会出现看起来令人讨厌的图案。因此，这些打印机使用称为“抖动”的另一种方法来表示图像。在简单的随机抖动中，打印机可能放置黑点的每个位置对应于图像中其灰度值已被检测的点处。如果该点非常暗，那么其打印的概率就较高，反之亦然。生成一个随机数并与灰度相比；如果该数低于该点（均适当缩放）的灰度值，则在页面上打印该点。

如图 3.8 所示，即使粗糙的抖动也可以产生可视的图像表示。使用了许多不同类型的抖动图案，其中的一些是随机或伪随机的，而另一些具有不同图案的点排列。通常情况下，所用的特定抖动图案由打印机本身或为其提供的接口软件确定。在抖动图案的设计方面，存在相当多的文献，但图像分析系统的用户可能没法较好地控制或选择它们。由于人类视觉对平均点密度或间隔产生响应，因此抖动的图像一般都能很好地显示了逐渐变化的亮度梯度。但锐利的边缘是模糊的、破碎的或移位的，因为不可打印出连续的线或边界，且也难以比较图像中区域的亮度。抖动印刷通常被认为是生产粗糙硬拷贝的“低端”方法，并不适用于报告或出版物。



(a)



(b)

图 3.8　第 1 章中图 1.24 所示“女孩”图像的两种不同抖动图案：(a)随机点；(b)图案点

尽管存在上述限制，但使用半色调方法的单色图像打印及具有 1200dpi 分辨率的打印机足以应付许多类型的图像处理应用，且所生成的图像也适合于报告要求的质量。由于办公静电复印机并不能完全再现打印件，因此需要为每个报告副本制作原始打印件。

对于更高质量的工作，可以使用摄影记录。这一过程看起来像是一种逐帧回放，尤其是在计算机中采集和处理数字图像时。由于处理摄影图像、复制它们及为打印而制作逼真的半色调网屏的设备发展逐步完善，因此这通常是最有效的解决方案。在只需要少量副本的情形下，强行使用在计算机内直接将图像并入报告中的“高科技”解决方案，可能在时间或成本方面是不值的。在页面上配以照片或将照片作为单独的页面插入，仍然是一种可行的演示方法。随着印刷技术的不断发展，以及文档创建软件能很好地融入图像，这种平衡将继续朝直接由计算机打印的方向发展。

3.3　彩色印刷

能够生成彩色图像的打印机要复杂得多，但通常用做标准的计算机配件。使用放大镜研究本书中的彩色图像时，会发现为几种不同颜色的油墨创建半色调并将它们叠加在一起来生成打印彩色图像的方法。图像照排机通常使用由计算机中的分色软件（为青色、品红色、黄色和黑色油墨）生成的四色版来进行全彩色印刷。这种做法并不唯一，尤其是在广告中，因为在广告中会引入额外的“专色”来确保色彩的完美匹配。由于彩色印刷工艺中过于复杂，因此下一节只介绍对大多数网络或台式彩色打印机而言重要的内容。

要在计算机显示器或电视机上显示一幅彩色图像，可使用电子束在阴极射线管（CRT）中照射红色、绿色和蓝色荧光粉来实现。用于平板显示器、投影仪及笔记本电脑的大型等离子显示器和液晶显示器（LCD）使用不同的技术，但也产生红、绿和蓝三原色。还存在有由离散 LED 或甚至电灯泡阵列组成的大屏幕显示器（如用于户外广告和体育赛事）。每种方法都以不同的方式生成颜色，但所生成的颜色仍然是产生可被显示的整个色彩范围的红色、绿色和蓝色的视觉混合。

图 3.9　CIE 色度图，显示了典型计算机显示器和喷墨打印机的色域

这一范围称为设备的“色域”（Stone et al.，1988）。在发出红色、绿色和蓝色光（RGB）的显示器中，产生整个范围中人眼所能看到的颜色部分是可能的。打印的图像则更加受限。图 3.9 显示了在第 1 章中介绍的 CIE 色度图，它比较了用于计算机显示器和

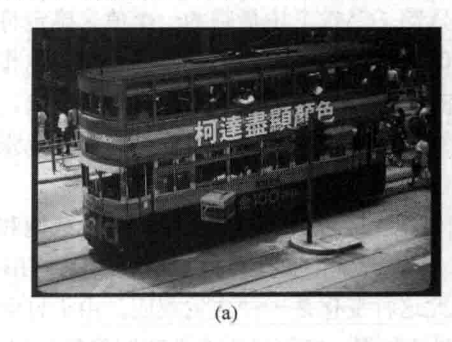
喷墨彩色打印机的典型色域。无法产生的其他颜色位于图形范围之外，它们是最饱和的颜色。

如图所示，色域这一概念适用于设备可以再现的颜色范围，但不适用于亮度值。这也是一个重要的问题，因为如第 2 章中所述，人类视觉能够对相当宽范围的亮度进行响应，该范围比任何显示器所能再现的都要大（且远远大于任何硬拷贝输出）。胶片相机和数码相机所记录图像的亮度值的范围，都要大于摄影打印机或油墨打印机所打印出图像的亮度值范围（S. Battiato et al.，2003）。这样宽的范围能为人类观看者呈现令人满意的场景表示吗？

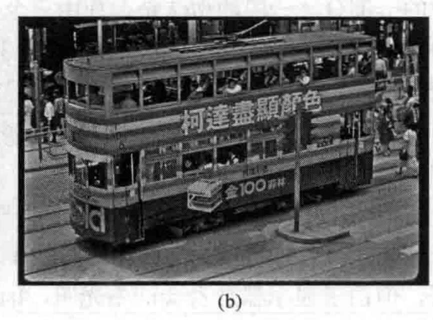
为达到此目标，人们在高动态范围图像处理方面提供了大量的文献［Debevec & Malik，1997；Durand & Dorsey 2002；以及过去十年间 SIGGRAPH（国际图形学年会）的论文集］。多数技术会在局部上保持相对亮度差异的同时，对亮度进行全局抑制，因为人类视觉会响应于亮度的局部变化，而往往忽略亮度的渐变。最简单的方法是一种高通滤波器（如随后章节中将介绍的滤波器，它会抑制图像中的低频或渐变，而保持或强调高频或突变）。这样的一些简单方法会在边缘四周引入光环，需要用更先进的技术才能将其消除。

有代表性的两种技术是 Retinex 方法和同态范围压缩。Retinex 方法基于 Edwin Land 于 1963 年的原创工作。如第 2 章所述，他认为人类视觉依赖于亮度和色彩的空间比较，这一思想导致了人们对视觉的许多深刻见解，即使是 50 年后的今天，这仍是人们的研究主题。2004 年 1 月出版的 *Journal of Electronic Imaging*（第 13 卷第 1 期）中，有一篇专题为“Retinex at 40”的回顾性文章，其中包含了用于亮度压缩的 Retinex 算法的一个 MATLAB 函数实现（Funt et al.，2004），当然也存在许多其他的文献（McCann，2004；Ciurea & Funt，2004；Sobol，2004；Cooper & Baqal，2004；Rahman et al.，2004）。

Retinex 方法在像素域中工作，并在不同的尺度进行比较，因此有时也被称为多尺度技术。同态方法则在傅里叶空间工作，它是压缩亮度值以匹配减小的打印色域的一种特别有效的方法。这一过程将在第 6 章中描述。如图 3.10 所示，结果在保持局部对比度的同时，压缩了图像中的整个亮度范围，以便可以进行打印。深阴影区域中出现了细节，同时保持了明亮区域的对比度。



(a)



(b)

图 3.10　亮度范围的同态降低揭示了阴影区域的细节，但未降低明亮区域的对比度：(a)原图；(b)结果

色域这一概念同样适用于其他彩色显示和录制设备。但色域并不总是如图 3.9 所示的形状那样简单。首先，印刷采用的是减色而不是加色。根据色彩立方体中（已在第 1 章中介绍并在图 3.11 中显示），白纸以白色开始，而青色、品红色和黄色油墨的相加会从反射光中删除互补色（分别为红色、绿色和蓝色）进而产生可能的颜色范围。我们通常使用这些减性 CMY 颜色来区分加性 RGB 颜色。

图 3.12 中总结的减色印刷理论表明，所有三种颜色混合在一起可以生成黑色（与加性 RGB 颜色一样，所有三种颜色的和产生白色）。但使用 CMY 油墨印刷并不能产生很好的黑色，而只能给出一种模糊的灰褐色，原因是油墨中存在杂质、油墨表面存在光反射（因此不能穿透所有将被吸收的油墨层），且很难完全覆盖整张白纸。

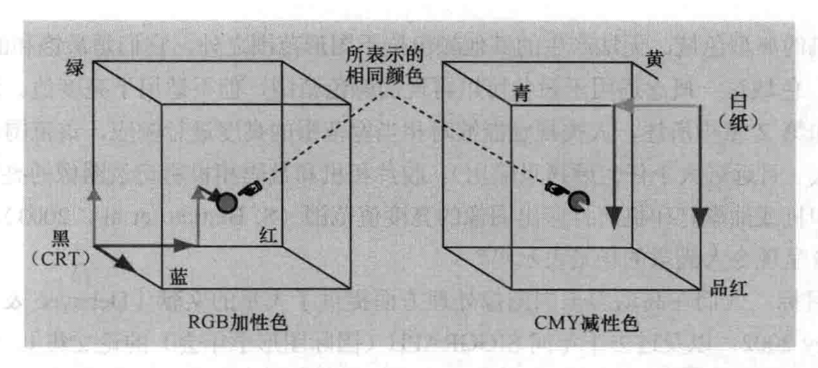


图 3.11 RGB（加性）和 CMY（减性）色彩空间的比较。加性色彩空间向黑色背景添加红、绿和蓝的发光色，而减性色彩空间从白色背景中移除青色、品红色和黄色

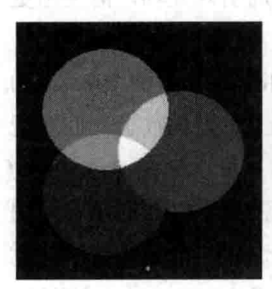
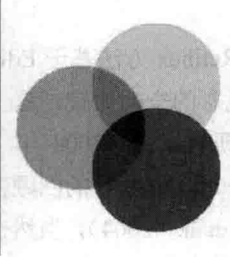



图 3.12 黑色背景上 RGB 加色的组合与白色背景上 CMY 减色的组合

最常见的解决方案是，在印刷工艺中加入一种单独的黑色油墨。这大大降低了彩色油墨的使用量，并减小了页面上油墨积聚的厚度，进而降低了成本。但从图像质量方面来看，最重要的因素是，这样做可以打印出暗色而不会出现模糊现象。青-品红-黄-黑四色系统（CMYK）使用 4 个半色调网屏，每个网屏对应一种油墨。然而，尽管从 RGB 色彩（或 CIELab、HSI 或任何其他色彩空间坐标系）转换到 CMY 很简单，但转换到 CMYK 并不容易。计算有多少黑色落入各种颜色表示中的规则，取决于我们对颜色的视觉响应、纸张类型、观看打印件的光照及图像中的内容（Sharma，2003）。

从 CMY 转换为 CMYK 的算法，会涉及底色移除（UCR）和灰色成分替换（GCR）的程度，这基本上是任意的，很少有文献记载，且不同软件的算法有很大的不同（Agfa，1992）。这些计算通常是打印机驱动软件的一部分。一般的做法是，使用三个分量（CMY）中最暗的一个值来确定将要添加的黑色量。例如，对于含有 80%青色、50%品红色和 30%黄色的颜色，30%蓝色这个值可以作为内置校准曲线或查找表的指标。这可能表明应选择黑色油墨的一个 13%的值来替代所有灰色成分。然后，不等量地减少所有颜色的量。很难为导致偏色的这些替换设计算法。此外，对于以油墨、纸张抛光度等的不同组合而言，替换并非能工作得同样好。

如图 3.9 中减小的色域所示，彩色打印件不如出现在显示器上的图像那样生动或饱和。此外，色彩严格依赖于纸张的颜色、抛光度及所观看的条件。室内光线的变化可能会稍微改变 RGB 显示器上颜色的视觉外观，但由于显示器自身会产生光照，因此这种变化是一种次生效应。由于打印件是通过反射光来观看的，所以改变打印件所在室内光线的光量或色温，就可以完全改变图像的外观。白炽灯泡、荧光灯、直射的阳光或天空的色温（描述亮度谱与颜色之间对比的一种简便方法）各不相同。日光的色温大约是 5500K，而白炽灯泡的色温通常是 2900K（更暖、更红），阴天色温可能高达 8000K（更冷、更蓝）。许多计算机 CRT 显示器的色温接近于 6500K；LCD 显示器的色温取决于采用的是荧光还是 LED 照明。

人类视觉可能会被光照和阴影、油墨或其他染色涂层、（光滑的或各种纹理的）表面抛光度的组合，以及视野中出现的其他相近颜色所欺骗，而这种欺骗会改变人们判断图像中颜色的方式。其他的颜色甚至可能位于图像本身外部，如彩色外框提升一幅艺术品外观的方式。需要一幅图像的“真实”彩色打印件时，必须对特定打印机和显示器进行校准，以便得到可接受的保真度。广告业中非常关注这一点；通过邮购商品目录购买服装的人期望服装的颜色与印刷照片的颜色相同，但这一任务并不轻松。

校准的方法是调整显示器输出（在显示器硬件中使用一张查找表），使得显示器的颜色外观与特定打印机/纸张/油墨组合重现的颜色匹配。

对于多数（但非全部）图像处理应用而言，彩色印刷的目的是区分所呈现的不同颜色的区域；色彩保真度中的某些不精确也许是可以接受的。如果在某个特定的应用中并不需要精确的颜色匹配，那么任务就会变得更加简单，尽管仍然需要关注打印机的色域、颜色的一致性（以允许比较不同图像或区域）以及打印机的分辨率。色域很重要，因为视频显示器上可以区分的原始图像中饱和度逐渐增加的颜色，在饱和度超出打印机范围的情况下，会在打印图像中变得类似。

尽管产生 CMYK 半色调（所谓的色彩分离）并将它们叠加来产生彩色打印件看上去很简单，但它同样具有许多商业打印机通常必须处理的问题。其中之一是如何叠加半色调网屏，以便不会在图像上产生视觉上的莫尔条纹。图 3.13 显示了由不同彩色图案产生的莫尔效应的一个例子。所示线条是各个图案间的间隔和对齐方式产生的人为缺陷。如图 3.14 所示，传统平版印刷对 4 个 CMYK 网屏分别定义了 45°、70°、90° 和 105° 的理想角度。这会对齐各个彩色网屏，进而形成在一起时构成颜色的小花环。图 3.15 显示了一个例子，网屏中的每个单元会改变用于控制每种颜色数量的角度，且有些半色调单元可能会部分叠加。由于一些油墨是部分透明的，因此可能会因打印顺序而产生不同的偏色。

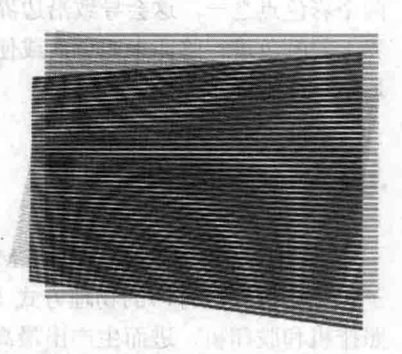


图 3.13　青色、品红色和黄色点交叠产生的莫尔条纹

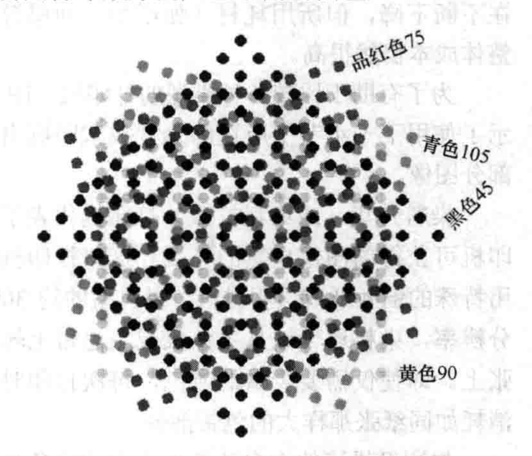


图 3.14　CMYK 彩色印刷的理想网屏角度会以 45°（黑色）、75°（品红色）、90°（黄色）和 105°（青色）的角度放置半色调网屏

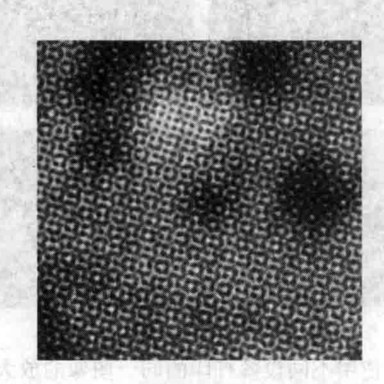


图 3.15　本书前一版中印刷的彩色半色调图像的放大例子

多数印刷并不需叠加各种颜色，而只需彼此相邻地打印，并允许显示一些白纸。多数彩色印刷方法会受到添加更多颜色、白纸因覆盖而使外观变暗的影响。这就是必须使用高分辨率半色调网屏和非常小的印刷点的另一个原因。一般来说，使用 4 个半色调网格或分色时，必须为每种颜色加倍打印机的 dpi 来得到与单色打印等效的打印分辨率。换言之，每个像素现在需要 4 个交错的半色调单元，每种颜色一个单元。由于每个 4 色点的大小只是该区域的四分之一，因此不可能展开打印某种颜色的实心区域，这会进一步降低可以达到的最大饱和度，进而降低打印机的色域。

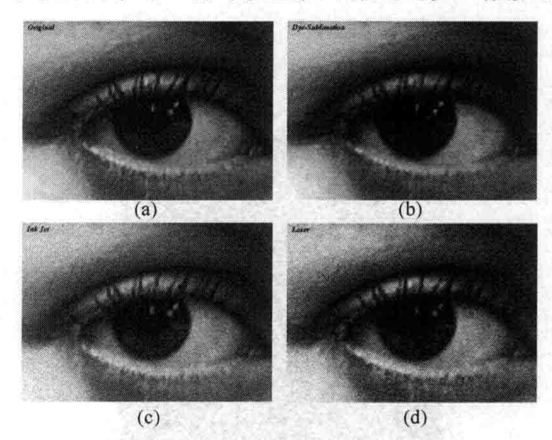
多数台式打印机不提供对网屏角度的控制，且由于打印方式的不同（如一种典型的彩色喷墨打印机），许多打印机会对所有 4 个网屏使用零角度。一些页面描述语言（如 PostScript Level 2）包含了对这种控制的规定，但通常只有高端的排版机和照排机能响应这些命令。

打印单色或灰度图像时出现的所有问题，如条带、色调分离、亮度级有限等，在产生分色时也会出现。此外，彩色印刷还会引入网屏角度和莫尔条纹等其他问题。最后，还必须考虑颜色的对齐和配准，因为许多打印机进行彩色打印时，是每次在纸张上打印一种颜色来处理的。有些打印方法会在一次淀积多种颜色，但这会出现油墨未充分混合就变干的问题。

彩色印刷和配准还会导致另一个问题。考虑一个颜色均匀的区域，该区域混合有两种或多种基本印刷颜色，且毗邻于一个不同颜色的区域。在边界，颜色一定是从一种颜色到另一种颜色的突变。但这两个彩色网屏并不会逐点对齐，沿边界可能会存在这样一种情况，即左侧的两个彩色点之一会靠近右侧的两个彩色点之一。这会导致沿边界出现假彩色线。在肖像画中，这些人为缺陷通常可以按某种顺序对边界上的颜色进行陷印来避免，或使用黑线掩盖住边界来避免。显然，这种方法并不能用于打印每个像素均不相同的真实图像。

3.4 印刷硬件

将图像转换为 CMY 值或 CMYK 值并在打印页面上将它们叠加为单独的半色调网屏的描述，忽略了执行这种彩色打印的物理方式（Kang，1997，1999）。很少有计算机用户能以足够高的分辨率来访问照排机和胶印机，进而生产出最高质量的结果。不同台式打印机的质量、成本和性能会有很大的不同。目前最常见的有效方法包括喷墨、染料升华和彩色激光打印机。随着众多厂商对庞大且不断增长的打印市场的争夺，打印质量在不断地进步。众多的技术指标中，成本最为重要。许多打印机本身的价格在不断下降，但所用耗材（如纸张、油墨等）使得整体成本仍然很高。



(a) (b) (c) (d)

图 3.16　使用不同设备打印的同一图像的放大片段：(a)原始图像；(b)染料升化；(c)喷墨；(d)激光

为了有助于比较各种类型的打印机，图 3.16 显示了使用下一小节中所述的几种技术印刷出来的一部分图像。

染料升华（或染料扩散）打印机代表了台式打印机可获得的图像质量的高端。这些打印机需要使用特殊的铜版纸或透明材料。以典型的约 300dpi 的分辨率，染料逐个地从主色板或主色带上转移到纸张上。即使仅需要少量的颜色，每次打印时也都要消耗如同纸张那样大的色带部分。

与这里描述的大多数其他类型的彩色打印机的 1200dpi 以上分辨率而言，300dpi 的低分辨率值并不表示打印质量，因为打印件是以不同的方式形成的。这些打印机可以更直接地控制每种颜色的数量，而不是装配小的打印机点来构建色彩。从色带转移到硬拷贝上每个像素的染料数量，是通过加热打印头升华或蒸发不同数量的染料来控制的。典型的控制是 1/256，为每种颜色使用校正伽马函数来平衡颜色，以便打印件看上去与网屏图像类似（所谓的“颜色匹配”），可能会将每种颜色的有效动态范围降低到大约 1/100。尽管如此，这仍意味着在图像的每一点上可淀积约 100 万（100^3）种不同的颜色组合。

染料升华打印机不使用半色调或抖动模式。染料扩散到纸张的聚酯涂层中并混合，以产生连续的彩色。纸张上每英寸 300 个打印区域中的每一个区域，包含了 3 或 4 种染料颜色的均匀混合。缺少任何半色调或抖动模式时，若在涂层内混合和扩散彩色染染，则会产生平滑且外观令人愉悦的图像。局部放大图像［见图 3.16(b)］显示了连续色调和恰好可见的像素图案。

有时一些打印机会声明输出质量“接近于照片质量”，但在几个重要方面，事实并非如此。由于饱

和度小于照片，因此色域较小，而以 300dpi 的分辨率在涂层上扩散染料进行混合，所导致的清晰度要比照片差得多（在照片上每英寸可很容易地分辨出几千个点）。沿纸张运动的方向而非纵向扩散时，结果会更糟，因为这会使得图像出现一些失真，或会使得图像的某些边缘更为模糊。在染料升华打印方面，大面积均匀或渐变的颜色看上去更平滑，因为没有抖动或半色调来断开它们，但边缘和线条可能显得模糊。

除了这些技术上的限制外，染料升华打印机及其耗料相对较高的成本，也限制了其用做各成像系统的专用打印机。同时，相当慢的速度（每次打印需要使三种或四种不同颜色的色带）也限制了其网络应用或制作多个副本方面的应用。尽管如此，它们仍是直接制作图像硬拷贝输出的首选方法。它们正在寻找一些商业网点用途，即消费者可以将幻灯片或彩色负片插入扫描仪，或从记忆棒或 CD 中载入图像，选择图像并在需要时屏蔽它，然后直接打印出硬拷贝带回家。铜版纸的光泽也许会使得许多消费者认为他们拥有了一张传统的照片打印件，尽管温度、臭氧和湿度对使用寿命的影响更糟。

这时，喷墨打印机就成为了广泛用于生成图像硬拷贝的设备。这种打印机的价格便宜（尽管耗材昂贵），且分辨率高（目前许多型号提供 1200dpi 以上的分辨率）。目前为止描述的所有技术中，喷墨打印机对纸张的质量最为敏感。高质量的打印件需要纸张具有浆内施胶和表面涂层，以防止油墨通过毛细作用进入纸张纤维（见图 3.17），同时提升附着力并提供明亮的白色背景。诸如 $CaCO_3$（碳酸钙）或 TiO_2（二氧化钛）这样的颜料也有助于使纸张更白。诸如明胶或淀粉这样的表面涂层因具有疏水性而可防止油墨扩张，也比未加工的纤维更光滑，还可以被滚轧（用压光机处理）以提供更均匀的表面。油墨必须使用非常微小的色素微粒精心制造，以便使得小液滴（目前最先进的是每个微小的打印机点为 1～4 皮升）均匀且一致地传输到纸张上。

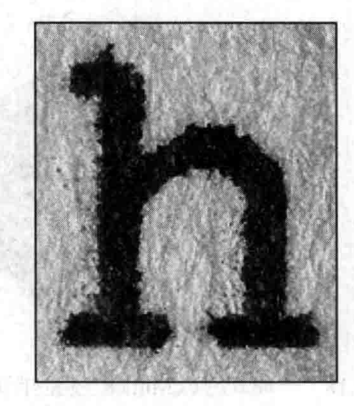
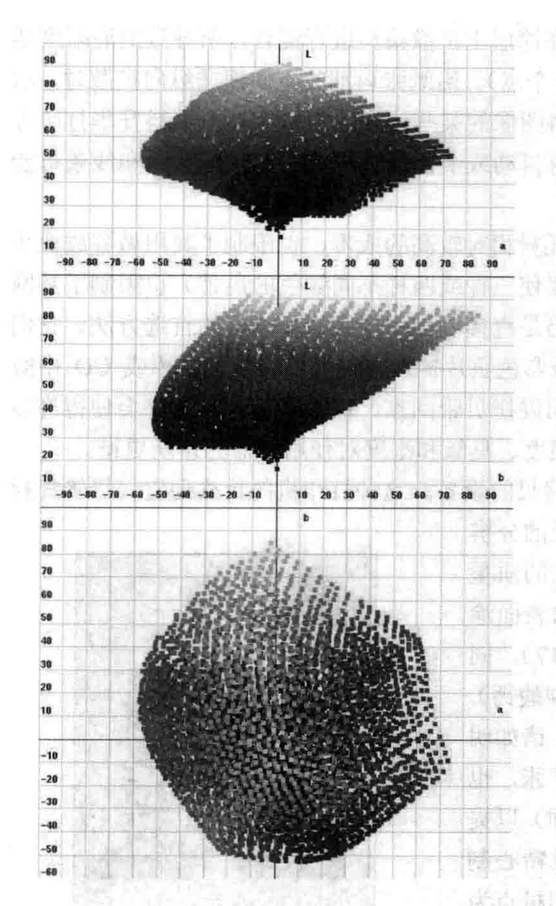


图 3.17　一张无涂层纸上墨汁的毛细作用

打印件的寿命会受紫外线、湿度、温度、臭氧、纸张酸度等的强烈影响。一些制造商现在提供“档案”品质的油墨和纸张，这意味着纸张具有较高的 pH 值（通常由添加的碱性缓冲剂来降低引起纸张变色或物理性能退化的酸度）和较高的破布浆配比（纤维而不是木纤维素），且油墨中含有不因紫外线照射而分解的色素。然而，除非存储条件能得到仔细控制，否则不要指望这些打印件能保存很长的时间。使得计算机打印的硬拷贝达到可接受寿命，一直是设备制造商所面临的挑战。提升打印件寿命的最新方法之一是，在彩色油墨的顶部“打印”一层附加的透明材料作为密封剂，以便防止氧化和潮湿。这也会增强最终打印件的视觉外观，并用于廉价的“照片打印机”中。

这些打印机中使用了两种类型的油墨，即颜料颗粒和染料。含有颜料颗粒的油墨受纸张变化（平整度、表面涂层等）的影响较小，且很少受光照、水分、臭氧或其他空气污染的影响。因此，它们通常被称为“档案”油墨，尽管打印件的实际寿命和对颜色变化的抵抗能力取决于存储条件。由于纸张或其表面涂层并不吸收颜料颗粒，因此颜料颗粒对涂污或刮痕更加敏感。通常，基于染料的油墨干燥得很快，且其反射性要强于基于颗粒的油墨，因此可产生更明亮的色彩和更丰富的颜色范围。

为扩充打印机的色域并提升渲染诸如肤色这样的精细色调的能力，许多喷墨打印机现在使用 4 种以上的基本 CMYK 油墨。在称为 CcMmYK 的典型六色系统中，也有两个较低饱和度的青色和品红色油墨可用。八色系统引入了较低饱和度的黄色和黑色油墨，其他系统可能还使用了其他油墨。对每个像素转换所存储的 RGB 值来恰当地混合多种油墨的算法是复杂的、独有的，并与纸张特性密切相关。许多喷墨打印机提供了多种质量等级。较快的设置所消耗的油墨较少（使用较低饱和度的色彩产生可接受的草案图像），而较慢的设置则用于最终的副本打印。


图 3.18 一种六色 CcMmYK 喷墨打印机的三维色域，它显示为穿过 Lab 色彩空间的截面（Rezanaka & Eschbach，1996）

在打印输出的较亮区域，可以看到单个液滴［见图 3.16(c)］。理想情况下，这些液滴都是同样大小且隔开的理想彩色圆点。实际中，它们可能会随着油墨喷嘴堵塞或磨损而迅速恶化，并受温度、吸收性和正在使用纸张的平滑度的影响。在要呈现更多的色彩并使用更多不同的油墨颜色时，它们会互相叠加并覆盖更多的纸张部分，进而产生更暗的图像。如上所述，这和 CRT 显示器上的图像的外观是相反的，在 CRT 显示器上，更多的颜色会产生更亮的像素。

如上所示，一台打印机的色域通常由可实现的最大色彩饱和度的色轮或 CIE 图上的轮廓描述。因为饱和度随亮度变化，故实际情况要复杂得多。由所打印出的覆盖整个三维色彩空间的色板，有可能构造打印机色域的一种更完整的表示。图 3.18 显示了一个例子，对于爱普生六色喷墨打印机，使用前面所述的彩色空间的 Lab 表示法。使用多种油墨的打印机可以生成数千种不同的色调，并且可以调节着色剂密度来产生高色调分辨率及颜色上的细微差异。然而，如所有这种色域图一样，仍存在于硬拷贝上无法表现的颜色。

喷墨打印机也可能配备有一种可选的灰色油墨，当用适当的软件进行控制时，这可以打印出非常好的灰度图像。通过选择不同的油墨密度并控制小墨滴的布局，可以构造与相片非常类似的灰度，且空间分辨率满足要求。选择具有良好表面抛光度和洁白度的纸张仍然很重要。

总之，喷墨打印机基本能够提供与传统感光材料的色调范围和分辨率差不多的硬拷贝（某些情况下所提供的硬拷贝甚至更好）。然而，尽管初始成本是适度的，但每份拷贝的成本与摄影打印件相当，且在许多情况下其使用寿命更值得关心。若使用得当（即针对所用油墨和纸张使用经适当校正的 ICC 曲线），那么喷墨打印机足以产生非常令人满意的结果，并仍然是众多用户的首选。

对于许多应用而言，彩色激光打印机是最经济的，因为它们使用普通纸张并具有相当高的速度。打印机的较高初始成本，可由每次非常低的打印成本来弥补。彩色激光打印机通过将图像分离成 CMYK 颜色，并使用每种颜色来控制一个单独的带电硒鼓的光照来进行工作，而带电硒鼓选择对应的彩色碳粉并将其淀积在纸上。这种打印机的一个主要设计问题是，保持所有东西对齐，使得纸张从一个硒鼓通过加热器烘烤碳粉并继续行进到下一个硒鼓时，不出现收缩或移位问题。使用具有平滑抛光度且相对较厚的纸张有助于解决这一问题。此外，所用色粉的顺序也至关重要，因为对于现有的多数型号而言，点并不是作为一组半色调网屏在不同的角度滴印的（如上所示）。相反，它们容易相互重叠，因而难以实现良好的色彩饱和度或准确的色彩调配。

图 3.16(d)表明，诸如红色这样的饱和色彩是通过两种或两种以上的碳粉色组合而成的，且能较好地覆盖白纸，但更精细的色彩不是那么清晰或饱和的，且可能会显得模糊。这类打印机的色域（颜色的最大饱和度极限）不是太有意义，因为最饱和的色彩定义得相当好，而较低饱和度的色彩却并非如此，且更难以彼此区分。尽管可能符合报表的要求，但彩色激光打印件通常仍是再现图像的一种边际选择。

也使用一些干墨打印机（由于油墨被溶解并转移到纸张上，所以也称为相变打印机）和热蜡打印机来打印图像，但它们更适用于计算机图形学，因为它具有鲜艳的色彩和光泽。这类打印机最适合于处理投影胶片，因此商业图形演示应用中使用得较多，但投影胶片对于图像处理输出而言，通常不是一种可接受的媒质。热蜡打印机的打印头中包含有加热元件，它会将蜡融化并将颜色转移到纸上。由于需要对每一种颜色重复这一过程，因此可能会引起对齐的问题。许多情况下，点太大，而每种颜色所生成的灰度数又太少，且分辨率太差，因此不能很好地显示图像的细节。图像呈现为压在纸张表面上而非纸张内的有光泽的平坦图版。

3.5 胶片记录器

为使用传统的胶片方法创建幻灯片或打印件，人们有时会尝试拍摄计算机显示器上的图像。除了在这一数字化时代获得所处理胶片的问题和延时外，这种方法还有一些潜在的问题，包括避免窗户或室内光在显示器上反射、CRT 显示器曲率所引起的失真（即使使用平板 LCD，因广角镜头引入失真而要求必须在一定远的距离拍摄照片），以及需要使用足够长的曝光时间（至少 1/2 秒）来捕捉多个光栅扫描并从局部扫描中消除不连续。由于显示器上的图像由彩色点的正常图案组成，因此该图案也被记录在照片上。如果照片随后被转换为半色调，那么从网屏上拍摄的点阵与半色调网屏的间距相互作用，就可能会产生令人讨厌的莫尔条纹。

胶片具有高分辨率、好的色彩饱和度及动态范围，因此是一种极佳的记录媒质。但拍摄彩色显示器不是记录图像的正确方式。专用胶片记录器通过使用具有连续荧光粉和一个平面的单色 CRT 来解决上面提到的那些问题。连续荧光粉给出了非常高的分辨率，并且不存在导致莫尔条纹的结构，而平面消除了失真。只用一种单一的电子束，就可以比 CRT 显示器更好地保持从中心到边缘的清晰聚焦和均衡强度，固定光学器件也可以无虚光或失真地提供横跨图像的清晰聚焦。

一种含有红色、绿色和蓝色滤镜的机动滤光轮放置在 CRT 与相机之间。类似于用来驱动一台打印机的软件被用来将图像分隔成红色、绿色和蓝色分量，这些分量在 CRT 上合适的位置使用恰当的滤镜每次一个地显示。更高的显示分辨率允许整个图像以全部像素分辨率来显示，这在计算机显示器上通常是不可能的。滤镜的密度被调整来平衡胶片的灵敏度。所有这三个颜色分量都记录在胶片的同一帧上。对大多数记录器来说，整个过程是自动的，包括胶片的传输和对不同胶片类型及速度的补偿。

唯一的缺点是几秒钟的曝光时间和直至胶片显影的延时，当然还有整个设备的成本。成本除与分辨率（从水平的 2000 点到高达 16000 点）相关外，还与处理不同尺寸和类型胶片的灵活性相关。高端记录器逐帧地曝光电影胶片，来产生包含“虚拟现实”场景及计算机图像的电影（如《玩具总动员》），或通过清除污垢及进行色彩校正来修复旧的经典电影（如《白雪公主》）。随着数字技术的不断发展，电影业正转而采用直接将数字电影交付到电影院来取代将胶片投影到屏幕上。

由计算机图像生成照片打印件的另一种方法是采用数量越来越多的在线服务。摄影工作室如今提供并不昂贵的从上传图像文件到打印件或幻灯片的转换。大多数通过邮件或联邦快递在几天之内返回打印结果。这是维护昂贵设备的一种可行替代方案，但大多数服务只接受 JPEG 压缩文件，且色彩保真度难以控制。处理图像文件的技术被用来处理真实生活场景的传统照片，而即使再进行颜色的调整也不可能在不同实验室或在不同时刻保持一致。颜色信息（如包含在由 Photoshop 产生的文件中的色彩空间描述）经常被忽略。面对科学领域或其他领域不熟悉的图像，这些工作室的转换结果通常会存在偏色现象。花时间按第 1 章中的说明来校准照相机、显示器和打印机的颜色后，令人沮丧的是，从这些工作室得到的结果是各有不同的。

3.6 其他演示工具

除产生图像的硬拷贝以包括在报告中外，图像的另一种重要用途是在 PowerPoint 和其他类似的演示工具中包含它们。粗心的用户会出现如下问题。

首先，许多图像分析程序可以在计算机显示器上以显示每个像素的方式来显示图像。有着百万像素的图像，为显示所有重要的细节，必须百分之百地显示出来，但其很可能会因大于显示器而需要滚动图像。跳过大多数存储的像素可以减小图像来适应显示器，但这可能会隐藏重要的细节。许多图像处理程序还会在图像文件上记录整个图像预期的输出大小（以英寸为单位），或记录图像将被打印的“dpi”（即像素数/英寸）。这些记录的内容接下来就用于控制打印机的输出，以便得到合适的结果。

当从该程序复制图像并将图像转移到诸如 PowerPoint 这样的演示工具时，所有这些信息及色彩校正参数很可能会丢失。在“幻灯片”上很难建立正确的图像大小。除非使用图像的一小部分，否则就必须降低图像的尺寸来匹配幻灯片内的区域，这意味着图像将被下采样到所需的尺寸，因而会丢弃许多像素。因此，很多细节可能无法在演示中看到，除非选择那些精心挑选的感兴趣区域。

其次，通过校准相机、CRT 显示器和打印机（在第 1 章中所描述的）而精心获得的色彩平衡和保真度，得到了为这些设备所存储的 ICC 曲线。但大多数版本的 PowerPoint 对其幻灯片不使用这些曲线，因此颜色在外观上与在图像分析程序中的是不同的。除非也校准投影仪，否则它不会匹配图像的屏幕外观。

如果图像被包含到设计用来经由网络通过 HTML 文档交付的演示中，同样重要的是要记住，网络浏览器一般不遵守或使用 ICC 曲线或任何其他形式的色彩控制。标准网络浏览器（仅）用 216 种颜色被认为是网络浏览器的标准色，但即使是特定用户显示器上的任何外观也很可能变化很大。以 pdf（Adobe Acrobat）文件分发文档是一种更好的选择，因为它确实提供了色彩信息方面的一些控制及所允许的图像压缩程度，但你仍然不能依赖于他人的显示器或打印机的校准来查看你的图像。

3.7 文件存储

当在基于计算机的一个图像分析系统中操作时，图像可以保存为磁盘文件。似乎没有理由认为这些文件与可能包含文本、图形或程序的其他磁盘文件有什么不同。在某种意义上说，就是这么回事：文件包含字节的一个集合，它可以表示任何东西，当然也包括图像。但是从实际角度来看，几个原因会使得我们处理图像文件的方式稍有不同，因为图像文件需要有不同的存储考虑：

（1）图像文件通常很大。对于计算机存储而言，一图胜千言的老话显然是一种极大的低估。单个黑白视频帧（通常以 640×480 像素存储）占用约 300KB 的空间，而全彩色的一帧大约需要 1MB 空间。对于目前具有超过 1000 万像素的检测器和对每个红色、绿色和蓝色各需要 2 个字节位深的数码静物照相机，文件大小可以是 60MB 或更大。台式扫描仪甚至会产生更大的文件。一组形成时间序列的图像或形成体素的三维阵列的图像（可被视为一组平面）会更大。一个 1000×1000×1000 体素的层析重建需要 1GB 空间，若体素的密度值具有超过 256 级的动态范围，那么所需的空间还要翻一倍。这意味着，所用到的存储容量必须很大，且最好是开放式的和相当快速的，例如允许使用某种可移动介质。因此人们需要在存储文件时采用压缩方法来减小文件的大小，且需要将降低分辨率后的文件副本（如缩略图）保存为可快速访问的索引。

（2）成像通常需要保存大量的文件。这强化了对大量快速存储的需求，但同时也意味着需要访问所存储的图像。构建一个可以用多种方式（包括向用户显示缩略图、使用关键词和其他索引工具等）访问图像的数据库，是一种已被许多软件生产商认可的重要要求。从图像中自动提取分类数据来协助检索是一个更加困难的问题。

（3）计算机时代的数据管理还没有充分探索复制那些包括文字和数字的相对简单记录的可能性。“文件柜”的比喻并不足以表现这一点。因特网上的搜索引擎编撰了在数亿万网站上使用的词语清单，以便能很快产生结果，但任何使用过它们的人都知道，该过程会找到许多无关的网站，但会错过许多人们感兴趣的网站。例如，谷歌提供了返回图像的搜索引擎能力，但检索根据的是相关的关键词，而不是图像内容。[④]

许多广泛使用的程序，例如 Adobe Bridge，允许从一个预先建立的清单中添加关键词到图像中，而使得使用这些关键词可相对容易地从磁盘上的一组图像中来查找图像。此时，用户仍须通过肉眼观察来选择感兴趣的图像。同时，这些程序并不能找到存储这些图像的正确 CD 或 DVD。

计算机文件存储的真正问题是访问文件并找到文档。将一个容纳一幅图像的域添加到主要用于构建容纳文本条目的数据库中，并不会把该图像添加到图像数据库中。例如，在人事档案中保存每名雇员的照片可能是值得的，但这很难让用户根据雇员的外貌或另一张照片来查找雇员。这时，需要查看档案中的每一张照片。

涉及图像的一类数据库本质上是反方向工作的。地理信息系统（GIS）存储有多幅图像和地图。这些图像和地图记录了不同类型的信息，每类信息对于定位都非常关键。还可能有文本记录与这些定位相关联。通过叠加和相关多幅地图表示，就可以将一幅地图上的位置特性（如道路或建筑物）与其他地图（各种类型的土壤或植被、地下含水层等）的特征进行比较。规划 GIS 时存在许多重要的问题，因为不同的地图和图像（包括航空和卫星照片）必须对齐和配准，并要考虑它们的不同分辨率。此外，显示如此丰富的多维数据，存在许多不同于大多数图像数据库的挑战。

设计图像存储最优系统时，可能还存在其他一些因素。例如，在取证、医学及其他应用中，保存对图像的所有访问、所有可能被应用的处理或编辑步骤的记录通常是重要的。有些程序将这些信息作为图像存储在同一个文件中，或作为伴随有实际像素值的“元数据”的一部分存储。

此外，如果图像将在用户之间或不同系统之间共享，那么存储格式的标准化也很重要。有些用户可能希望将图像永远存档，但很少访问任何特定的图像，而另一些用户可能希望不断并反复访问这些图像，但希望在一定时间后或当项目完成后丢弃图像（或至少将它们移动到闲置的存储空间）。数据可通过网络供许多用户访问，或只限于单个用户访问。如果一组图像将要作为永久和不可改变的资源广泛传播来用于参考用途，则需求会再次改变。

已有一些解决这些问题的方案，且迹象表明会出现更多的解决方案。例如，我们很容易就会想到有人会试图寻找“Sarah 阿姨”的旧照。如果照片以褪色的打印件保存在一个旧鞋盒中，那么搜寻者会有在阁楼中来寻找鞋盒的一个初步想法，然后整理内容来寻找图片。但在计算机时代，我们自然会要求计算机来寻找该照片。显然，这一任务并不简单，但许多公司正致力于应对这一挑战。甚至苹果公司的 iPhoto 程序也像一些数码照相机那样，包含有在快照上自动识别面部并进行标记的基本工具。

3.8 存储介质

图像可以许多不同的格式录制在 CD 和 DVD 上。考古发掘、医学研究和遥感（卫星成像）等项目表明，以 CD 或 DVD 形式来分发图像非常方便和廉价。CD（光盘）或 DVD（数字视频光盘，包括蓝光高容量 DVD）只是一种在其上压印有反射点的塑料盘片，其物理上与音频和视频光盘相同。CD 上点的格式标记从 1 至 0 的变化，而不是实际的比特值；编码模式最大限度地减少了存储原始数据所需要的比特数。

④ 在本书翻译过程中，已经出现了以图搜图的搜索引擎，如百度识图（http://stu.baidu.com/）、搜狗识图（http://pic.sogou.com/shitu/）及谷歌图片（按图片搜索）等。——译者注

CD 或 DVD 上压印有大量的反射点，为防止其毁坏，反射点上覆盖有耐磨损的硬质塑料。这种技术使得廉价地创建多个磁盘副本成为可能。由母盘制作一个副本的价格目前约为几十美分，且价格正呈下降趋势。如发布软件那样，以这种方式发布图像数据库很有意义，因为实际上每台现代计算机都有一个这类磁盘的阅读器。

CD 和 DVD 上的数据按一个连续的螺旋轨道写入，每块的大小为 2048 字节（加上块地址和一些检错和）。整个轨道被一次写入，而数据格式并不设计为用于写入或读取的一个随机存取装置，就像有多个同心圆磁道和固定扇区的传统磁盘驱动器一样，这使得对特定的数据块进行定位更加容易。

由母盘压制磁盘会在涂有铝的塑料基板上创建许多小坑，然后用一个保护性塑料层进行密封，这一塑料层不仅可提供物理上的保护，而且可阻止会降低其反射率的铝的氧化。在可刻录光盘中，写入是通过激光完成的，激光是在染料层中被吸收的。染料分解并淀积在金属层上，使其变暗，并产生与在压制磁盘上的物理凸面和凹面相同模式的反射区和暗区。

可写 CD、DVD 和蓝光光盘已被广泛用做存档文件的“一次性写入”存储设备。具有读写能力的 CD 和 DVD 驱动器是大多数台式计算机和笔记本电脑的标配，而在本书写作时，蓝光驱动器也变得越来越普遍。文件被写入到一张空白的盘（其成本只有几个美分）上。写入过程相当快；通常情况下，只需要几分钟就能写满一张盘。

CD 提供了高达 700MB 的存储容量。在被写入后，CD-R 光盘可以在和压制的 CD-ROM 完全一样的驱动器中读入，或作为母盘去创建大量更持久的压制光盘。由于 CD-R 光盘已变得非常可靠和廉价，因此当物理磁盘比在因特网上传输更适合时，它们已经取代软磁盘来传输甚至少量的数据及分发计算机软件。因特网传输无法与在联邦快递信封中的一些 CD 或 DVD 的带宽相比。

使用同样大小但可擦写的可写 CD（CD-RW）要稍微昂贵一些，且丧失了防篡改的优点。可写但不可更改光盘尤其适用于那些必须永久保存图像的情形（如医学和取证成像）。这样的驱动器也适用于记账和其他需要保证数据不被篡改的应用。

DVD 的存储容量更大。单层磁盘的容量为 4.7GB，双层磁盘的容量为 8.5GB。蓝光驱动器使用 405nm 的激光（蓝色）进行读写。比 650nm（DVD）或 780nm（CD）红色激光更短的波长可增大数据密度，使得单层磁盘的容量高达 25GB。双层驱动器很常见，四层驱动器也见诸报道。

CD 和 DVD 的缺点之一是，多数情形下，写操作必须在单个会话中进行。主计算机上必须拥有将要放到在盘上的所有文件，且在同一操作中写入，以便可为后续的读取创建磁盘目录。在创建 Photo-CD 格式时，柯达引入了多目录（或多会话）的想法，使得光盘能收进照片修整器，进而在一张已有的光盘中添加更多的图像（如另一卷胶片）。然而，并非所有能够读取 CD 的驱动器或软件都能够访问附加的目录，它们作为单独的驱动器而非一个大容量的卷出现，且存在与每个额外会话的目录相关联的开销（约 9MB）。

CD、DVD 或蓝光光盘的实际用途之一是，为存储或分发生成整个磁盘映射或其他数据。图像本质上用于存档（尽管光盘尤其是蓝光光盘会因处理或存储而毁坏）；用户不能在原始光盘上修改它们。从 CD、DVD 和蓝光驱动器进行读取通常相当缓慢，这种读取速度与读取计算机的标准磁性硬盘的速度相比，要慢一个数量级。但对于存取存档的图像而言，这种速度是可以接受的。主要总是在于定位正确磁盘所需的时间很长。

3.9 磁性记录

计算机中最为通用的硬盘使用磁性存储。（玻璃或金属）盘片上涂有一层薄薄的磁性材料，就像磁带那样。读写操作的原理如下：高速旋转（速度为 4000～10000rpm）磁盘表面，使得上方的磁头因旋转产生的气压“掠过”而不接触表面。使用半导体存储器和没有移动零件的固态驱动器，存储

速度更快，但成本会更高，且与传统硬盘驱动器相比，容量会小一些。两种类型驱动器的存储容量仍在增长，但即使是万亿字节的硬盘存储器也会很快被图像填满，因此，备份需要一些可移动存储器或外部存储器。

容量为几百兆字节的可移动磁存储磁盘几年前就相当普遍（如广受欢迎的“Zip”盘），但现在已经被固态存储设备（最常见的是插入 USB 端口的闪存“棒”，本书写作时，其容量为 128GB 或更大，其成本使得它更适用于传输信息而非存档）所取代。我们需要可移动类型的存储器，因为任何传统硬盘，甚至容量为数百 GB 的硬盘，迟早会填满图像。任何旨在进行图像分析工作的系统，都需要某种形式的可移动而无限制的存储空间。

由于传统硬盘的访问速度很快，因此适用于将文件复制到硬盘上进行存取，仅在工作快要结束时才将文件写到可移动介质上。采用这种方法，硬盘的作用就是为正在进行的工作充当短期本地存储位置，但可移动媒介仍用于为长期目的存储图像。硬盘也是存放计算机操作系统、所有程序（包括图像分析程序）、这些程序所产生的临时文件的位置。

存储器的最新进展之一是带有冗余硬盘阵列（RAID 存储）的网络服务器，这些阵列并行工作，因此有着更高的速度和错误保护。通过以太网或 WiFi 传输大量图像要比局部存储更耗时间，但优点是将维护和备份存储的责任转嫁给了组织内的 IT 专家，且图像适用于网络上所有计算机的可能性，使得其在某些情形下更有吸引力。包括图像文件在内的信息“云”存储（即在网络上的某个地方、在某些未知的位置并由其他人维护）是可能的，但因为网络传输速度相对较慢（且文件比较大），并且因为存在对安全的担忧（损失、未经授权访问或篡改的风险），这尚未成为首选的解决方案。

与光学存储相比，硬盘存储的存档性要差一些，因为杂散电磁场或高温可能会导致数据的意外擦除，而且不断的使用会增大损坏的可能性。CD 或 DVD 声称最适用于存档，但这种说法只适用于压制的光盘；可刻录光盘声称在良好的存储条件下只有 5～10 年的寿命，且与压制的光盘相比，其表面更易于被毁坏（蓝光光盘更是如此）。由于技术的改进，未来人们更为关心的是 CD 或 DVD 光盘是否可读。

如果目的是图像的归档存储，那么无论选择何种介质，聪明的做法就是投资额外的驱动器和接口。10～20 年内，要求通用存储设备仍能读取当前的媒介并不明智，因此需要一整套备用配件、软件驱动程序和工作体系。对于接口和网络而言，同样如此：几年前的 SCSI 和并行接口现在已被 USB 和火线（Firewire）所取代，而以太网和 WiFi 已经取代了令牌环网和 Appletalk，在几年之内所有这些都可能被其他技术所取代。通过不同的接口和网络将图像传送到下一代电脑中是很困难的。如果怀疑这一点，可考虑 10～20 年前常见的存储介质（8 英寸软盘、磁性光盘、Bernoulli 和硬磁盘盒等），这些介质都不能被当前大多数计算机读取。当然，还存在一些极不寻常的情况，如 20～40 年前的打孔卡片和 DECtape 那样的存储器，现在已经不可读取了，因为设备都消失了或陈列在博物馆中。

归档存储的另一个要求是，要用适当的安全和防火手段维护一些远程站点的备份文件，并且具有在那里存储正常备份的常规方法。如果数据值得保存多年，那么在每个周五将文件备份到 CD 上并在回家的路上将备份放至保险箱中实际上是不够的。遥感磁带（顺便说一下，这是改变存储方式的一个较好例子——多数早期的磁带只能在少数得到心保存的驱动器上读取）、医疗记录和图像等的情况就是如此。对于许多种类的记录，特别是图像，感光胶卷仍然是供选泽的方法；它能保存数据几十年，占用的空间相对较小，且能使用标准设备存取。

但对于多数人来说，存储大量图像的目的并不完全是为存取需要而归档保存。找到以前的图像，与现有的图像相比较，以发现并测量相似性和差异的能力，在许多图像处理任务中显然是需要的。选择一种存储方法并不是满足这种需要的困难部分。选择上面提及的存储方法时，可以根据在成本、存取的频度和速度以及所需要存储空间的大小进行。技术挑战在于图像存储后的找到一幅特定的图像。

3.10 图像数据库

保存大量图像自然就提出了如何定位和检索特定图像的问题。若存储使用可移动介质，则问题会更复杂。任何大型数据库都存在这些问题，如人事或销售记录数据库。搜索图像时会涉及某些特殊的问题。有些数据库管理程序会带有一些专用于图像的功能，但有些程序还不具备这些功能。例如，Adobe公司的 Bridge 程序可显示每幅图像的缩略图，并且有可搜索的元数据（“关于数据的数据”，或随图像提供的描述性数字和文本）。除了如日期、时间和图像格式细节之类的系统信息外，有些数据可由用户指定（如幻灯片标识、操作员、放大率、注释、版权信息、关键词的选择等）。它还可以记录对图像执行的图像处理操作的全部细节，这对许多应用（如取证）是非常重要的。测量程序可以向记录中自动添加参数值（如所呈现的目标数目、测量结果或色彩信息），或访问它来提供其他的搜索或分类可能性。扩展元数据以包含识别访问该文件的人员的身份的信息，或加入诸如在显微镜载玻片上阅读条形码等其他类型的数据，在这些方法中是可以办到的，但尚未实现。

多数数据库管理程序都能基于关键词的逻辑组合或搜索字段的内容来搜索词条。例如，对于图像，从连接到一台特殊显微镜（使用透射光穿过一个 500nm 的彩色滤镜）的照相机中，可搜索 10 月 1 日和 15 日之间记录的（由对应于患者 ABCDE 的 12345 号幻灯片获得的）图像。也可以查找那些包含特殊特性的图像，例如某个范围内的卫星图像，这些图像的红外信号表明有一片丰收的藻类作物。

如果对于每幅图像已将分类数据输入到适当的数字和文本字段中，那么由已有的数据库搜索程序，可处理这些任务中的第一项任务。第二项任务要求强度更高的计算机处理来从每幅图像中提取所希望的信息，因为这些特定的信息一般不可能已预先确定并输入到搜索数值字段。预料到今后可能会问到什么问题，通常并不实际，但这是有可能的。对图像数据库回顾性地执行分析，需要紧密结合数据库管理程序和图像分析软件。对于特定应用（如子宫颈抹片检查幻灯片的自动扫描和筛检），可以实现这样一种系统，尽管已进行了大量的工作，但仍未出现通用且灵活的解决方案。

通过多个数值字段或关键词列表进行的搜索，通常是是由布尔逻辑规定的，如创建日期必须位于某个值之前，与另一个必定出现的特定关键词（而且多个关键词必须以较接近的方式出现在描述性文本中）进行“与”（AND）或“或”（OR）运算，但图像必须“不是”彩色的。这类搜索与其他类型的数据库管理程序和因特网搜索引擎中使用的那些搜索相似。更为有用的一种搜索采用模糊逻辑。例如，查找“圆形”和“黄色”的特征并不是只指定用来判断圆形的标准，也不是只指定所需的数值范围，更不是只指定表示黄色的颜色分量组合或色调值范围。

有时，“按图像内容查询”（QBIC）或“基于内容的图像检索”（CBIR）这种方法，会试图包含像图像目标和区域的颜色、纹理和形状这样的视觉上重要的规则（Pentland et al.，1994，1996；Flickner et al.，1995；Yoshitaka & Ichikawa，1999；Rui et al.，1999；Mitra & Acharya，2003；Tushabe & Wilkinson，2008；Datta et al.，2008）。取决于所用的字段，不同的规则对于搜索而言可能是重要的。例如，在某个医学应用中，用户可能希望找到“包含有这种纹理的肿瘤的其他图像”（这意味着目标为肿瘤，以及纹理的合适数字描述），而在监控中目标可能是“包含与这架飞机形状相似的其他图像”。这样的一个搜索的关键问题包括图像和目标属性的推导与计算，以提供有用的查询功能，以及与精确匹配相对的基于相似性的检索。开始时，可以只对一幅或多幅示例图像进行搜索（“实例查询”），之后对搜索进行优化。对于这样的查询系统，有效的人机界面也在发展中。

基于未指定图像内容（颜色、图案、特征数据等）的例子的搜索中，程序很难预测哪些内容是重

要的，因此"实例查询"是有很多缺陷的一种想法。考虑图 3.19 中的例子：左上角的图像作为一个例子，它显示了几个响应。但预期目标是什么？粉红色背景上的白色特征重要吗？（从一幅图像中自动地拾取颜色很容易）。如果不是这样，搜索将会排除很多图片。有乌黑眼睛和鸟嘴的大白凤头鹦鹉是一个很好的匹配吗？如果目标只是哺乳动物而不是鸟，则答案为否。白色的小猫长有长毛，像小狗，但拉布拉多（Labrador）是只狗而不是猫。也许图像外部的圆形很重要，或存在两种动物。所有这些假设将搜索限定到相关的动物图片数据库，否则搜索结果可能会是汽车、花卉等。

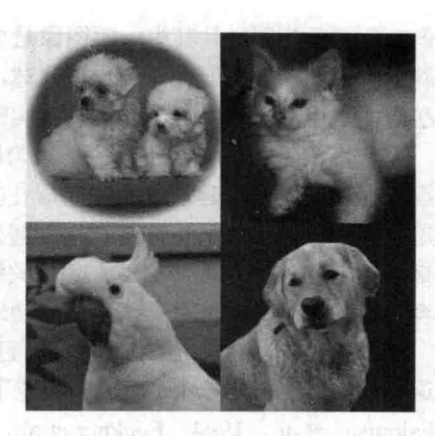


图 3.19　正文中讨论的"实例查询"结果

事实上，如果在无注释的情况下显示示例图像，而目标是找到其他类似的图像，那么会出现相同的不确定性。但在这种情形下，人们可能会询问有关图像的问题，以限定感兴趣的重要特征。

对"圆形"和"黄色"特征的例子进行搜索时，可能需要首先测量数据库中图像特征的代表性样本，才能得到圆形（或许是最短尺寸与最长尺寸之比，其他相关的形状规则将在第 11 章中描述）和色调的测量。例如，"圆形"可被用来表示位于最高 5%～10%的数值范围内的物体，而"黄色"是包括真黄色及包括观察结果的相似部分的色调值范围。

我们希望用于显示示例图像的程序然后可用于搜索所有的相似图像，但目前为止这类程序并不存在。部分原因在于很难事先知道图像的哪些方面在之后会是重要的。当人们搜索鞋盒中的照片时，所用的搜索标准通常会变化很大。我们通常能在视觉上区分这些照片，但搜索程序如何知道谁是 Sarah 阿姨呢？能因为她未去过海滩而忽略一幅海滩场景的照片吗？如果某人不知道背景，那么图像内容的意义通常会是模糊的。

例如，图 3.20 显示了在因特网上使用搜索引擎以关键词"ACC+篮球+比赛+照片"进行搜索时，在数秒之内（从总数超过 60 万的图像中）找到的一些图像。注意，其中的几幅图像中存在球员和篮球，

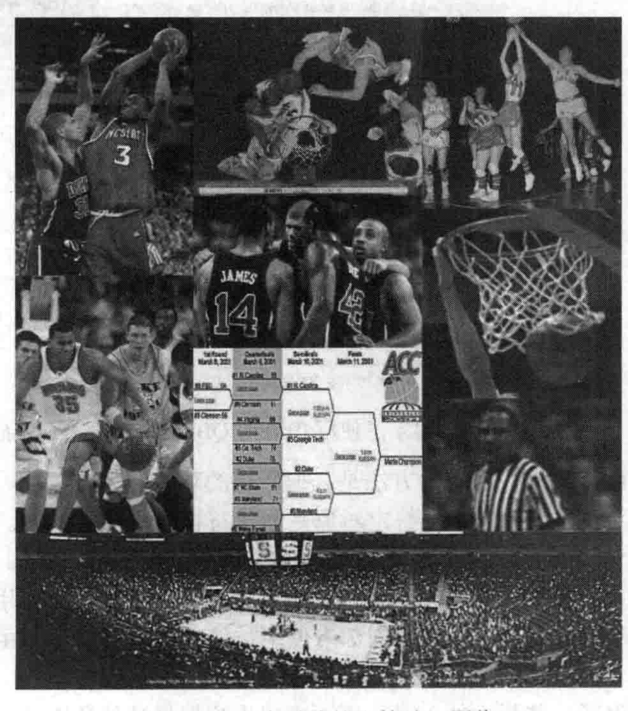


图 3.20　文中讨论的"篮球"图像

但颜色并不相同（且其中一幅图像是黑白的）。有些图像中不存在篮球，而另一些图像中只有篮球和球网。有三幅图像中既无球员也无篮球，但可识别为与篮球有关的其他事物（裁判员、赛程和篮球场）。在当前及可预见的发展态势下，向实例查询显示任何一张图片，都不会找到其他的图片。

许多科学成像情形下会出现更为困难的场景。可能存在大量看起来相似的图像，如患者肝脏活体组织切片图像，或建筑物密集的居住区的卫星图像，此时定位特定图像的视觉和机器方法几乎完全无用。有经验的操作员也许能识别一些感兴趣的关键特征，但不能在软件中用一种有意义或独特的方式来描述它们。在这种情况下，基于关键词、日期、病历、位置等的搜索可能更为合适。另一方面，内容查询系统已根据艺术家的风格成功地找到了绘画作品和其他艺术品（Holt & Hartwick，1994）。

几个商用程序声称能提供某种程度的图像搜索功能。最知名且文档最丰富的是IBM的QBIC系统，这是由IBM的Almaden研究中心花了15年时间开发的一种原型开发工具包（Niblack et al.，1993；Faloutsos et al.，1994；Flickner et al.，1995）。它允许用户基于颜色、纹理、形状、位置和布局来查询一组图像。例如，“查找一幅绿色背景且左上角有一个红色圆形物体的图像。”该系统及其他类似的系统首先会处理文件中的图像来提取诸如直方图（亮度和色调的均值和标准差）这样的信息，将图像分割成均匀的几个区域，测定其大小和形状，然后通过这些鲜明的特征进行搜索（Marsicoi et al.，1997）。一个有效系统的例子是联机搜索俄罗斯圣彼得堡冬宫博物馆的内容的能力。选择颜色，排列这些色块来表明它们的合适位置，可用于查找图库中的绘画作品（见图3.21）。

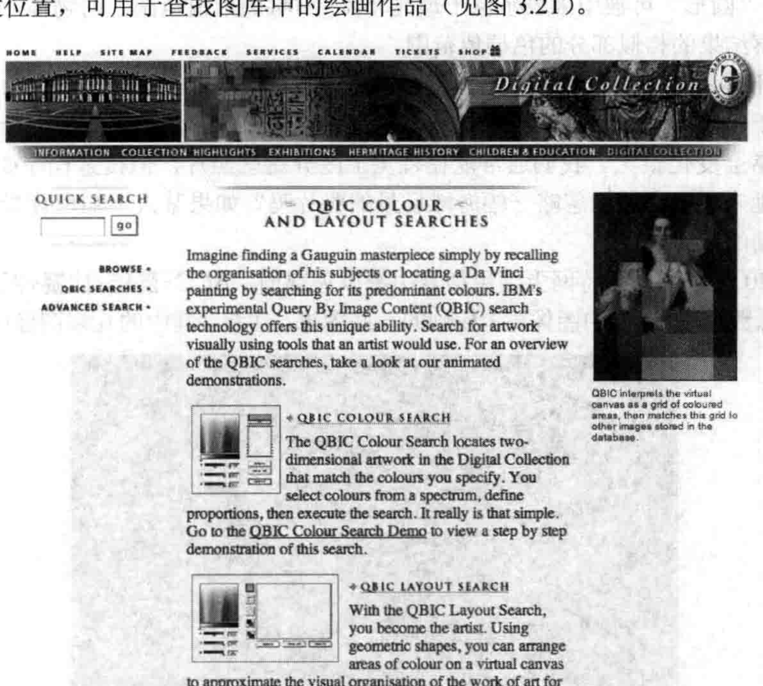


图3.21　冬宫博物馆网站显示了使用IBM的QBIC系统来搜索藏品的方法

将模糊条件解析为数值的一种强大的方法是，把所有的观察结果按尺寸、色调或其他一些客观标准进行排列，然后用排名序号来确定某个特定对象是如何符合所用形容词的。如果一个特征在几百个特征中，其圆度排名第9，黄色度排名第17，那么可认为它是圆形和黄色的。另一方面，如果它的一个属性排序第79，那么它可能就不是圆形和黄色的了。这两种方法最终都使用了数值，但依据特定类型应用中所出现的观察范围来判断结果。它们都未使用传统的参数统计，即由均值和标准差或其他描述子来表征一个群体。

模糊逻辑为分类提供了一些强大的工具，而不需要用户根据数值范围进行决定。但是，尽管模糊

逻辑已积极应用于控制电路和消费类产品中，但它在数据库搜索中的使用仍然相当有限。原因在于，与测量大量具有代表性的图像以在特定的上下文内给用户定义的含义所需要的时间相结合，已经延迟了这种方法在图像数据库管理中的使用。因此，大多数的搜索使用特定值（如创建日期在 2010 年 1 月 1 日后，而不是“最近”）和用户指定条件的布尔匹配来进行。

另一种搜索策略使用 N 维参数空间中的向量（Kang & Leon，2005）。多个测量参数的值被存储，且对于一幅给定的图像，之前由数据库程序索引的最近聚类，对 N 个最接近的匹配进行搜索。每个轴（或测量参数）可能有不同的权来确定 N 维距离。第 12 章将深入探讨这种类型的分类。

通常情况下，在数据库还未建立和存储图像之前，总由用户来建立对未来搜索有意义的标准、测量参数或诸如图像元数据（放大率、病历、日期等）的其他因素。如果还未设置适当的字段，就没有地方来记录这些值。高估预先设计的关键词或参数列表的正确搜索字段是不可能的，而且很难再次增加更多的字段。由于每个图像处理应用都有独特的要求，因此建立字段没有简单的指导方针。用许多关键词建立几十个甚至上百个搜索字段，对任何图像数据库管理程序来说都是最低要求。

当然，在保存图像时仍需要用适当的数值和关键词来填充搜索字段。有些字段可被自动填充，如日期和时间、操作者姓名，或诸如放大率、波长、位置（可以是经度和纬度值——许多照相机在图像文件中包含有 GPS 坐标——或显微镜载玻片上的坐标）等。甚至患者或样本 ID 号码也可以自动地登入；一些实验室使用条形码标签和阅读器来自动完成这一功能。在图像采集和保存时这样做并不烦琐；事后对一组图像提供这样的信息是困难且容易出错的。

但真正感兴趣的关键词和描述子仍然需要人类观察者填入到字段中。识别和辨识图像的内容、挑选重要的特性或特征而忽略其他，然后选择输入数据库的最有用关键词或数值，都高度依赖于操作者对图像和程序的操作水平和熟悉程度。

词条变化很大并不令人奇怪。不同的操作者或不同时间的同一操作者，可能会产生不同的分类。有助于限制操作者选择关键词的一种方法，通常是建立一个操作者可自其中进行选择的词汇表，而不是描述性文本的自由格式词条。该词汇表稍后用于选择搜索项的逻辑组合。

包含在数据库中用来测量参数的值，通常可以从图像分析程序中获得。后面关于图像测量的章节表明，特征的自动计数，大小、位置和密度的测量等，可由计算机软件来完成。但通常这要求人们来确定图像中的测量内容。这些程序可能会也可能不会直接将数值传递给数据库；在许多实际情形中，需要人工将值输入到数据库中，这就有可能出现疏漏和差错。

即使是在完全了解目标的情形下，如检查血液涂片，或扫描金属表面寻找含有一种荧光染料的裂缝，或在军事侦察照片中寻找交通工具，目标特性、样本准备或成像条件、照明等总会有足够的变化，因此需要人的监督来防止差错。而且通常在这些情况下，图像本身并不需要保存，只需要测量结果。当图像相似时，在数据库中保存所有图像是最常见的，但由一些数字或关键词并不能足以描述这些变化。

一个重要的取证目的是，能够使用一幅图像（可能是一幅指纹图像或由监控视频捕获的某人的脸部图像）进行自动识别，方法是找到一条包含有类似图像的计算机记录。这种技术的局限性之一是，示例图像的质量通常非常差。监控视频存在很多问题，如其广角镜头会使得面部很小，相机位置的光照条件较差，分辨率低，而且有时会使用质量较差的存储介质。犯罪现场得到的指纹与 FBI 数据库中的指纹相比，质量通常会很差。这种方法的另一个问题是，记录的比较基于对示例图像的测量，而这些测量难以自动进行，往往需要人的参与。最后，即使假定某人的记录在文件中，匹配也很难完美。目前，这些程序所能做到的是，找到最为匹配的最少记录，并将其提交给人来进行核对。

一类非常不同的数据库使用图像而非单词或数字来作为组织原理。数据库中的数值或文本数据对图像上的位置非常关键。如上所述，这种方法通常称为地理信息系统（GIS），通常在图像是基本上是地图时使用。我们可以想象这样一种情形：获取了同一区域的许多航空和卫星照片。有些图像可显示可见光信息，但也许有不同的分辨率，或是在一天的不同时间或一年的不同季节所拍摄。其他图像可

显示该区域的红外波段信息，或是显示高程的距离图像。还可能有显示道路和建筑物、土地使用模式及特定目标（消防栓、电力变压器）位置的地图。与所有这些相联系的可能是其他类型的数据，如邮件地址、电话号码和人名。

其他种类的信息可能包括温度、交通模式、作物产量、矿产资源等。将车辆沿路拍摄的地面图像与卫星图像相关联的谷歌“街景”（Street View），就是一个将不同类型数据汇集在一起的例子。

按位置组织这种信息并不简单。首先，各种图像和地图要配准后才能叠加。分辨率和比例尺通常是不同的，且信息的“深度”同样也会变化（每个像素随之由多个亮度值组成）。找到访问并呈现这些数据的方式，以便对不同类型的信息进行比较，对我们而言是非常大的挑战，甚至搜索数据库也没有我们想象的那么简单。

例如，选择坐标（如经度和纬度），然后要求在该位置的各种类型的信息相对简单。目前已出版计算机化地图（以 CD 发布或存在于几个网站上），通过连接到 GPS 接收机的便携计算机可对这种地图进行访问，而 GPS 接收机则拾取来自卫星网络的定时信号来计算出地球上的任何位置，精度可达几米。很明显，这可用于提取正确的地图部分。但当驾车使用这种系统来导航时，问题可能就会是“请显示到最近加油站或医院的路径，或者避免下午 5:00 的塞车并到达主要公路”。显然，这些问题需要访问有关街道、交通信号灯、时间和许多其他的附加信息，这就是包含 GPS 功能且可访问谷歌地图的智能电话如此有用的原因。将从网络获得的信息叠加到手机照相机拍摄的图像上，仅是融合丰富数据集的最后一步。这是目前的技术；未来的期望则是用更丰富的数据库和更灵活的方式来访问与显示信息。

3.11 浏览和缩略图

本节首先介绍如下想法：保存图像，以便它们以后可与其他图像进行比较。这种想法是完全必要的，因为没有简洁的方式来描述图像的所有内容，也无法预测到图像的哪些特性或特征在之后可能是重要的。这表明，在数据库中使用简短的描述子可能无法提供一个稍后足以定位特定图像的工具。因此，除了可以用到任何逻辑搜索条件外，大多数图像数据库提供了一种方式让用户来查看图像，或至少是一个其低分辨率的“缩略图”表示。尽管对有着相似之处的很多图像的缩略图进行视觉搜索来找到感兴趣的特定图片明显是困难的，但这仍然是最广泛使用的方法（Hogan，2005）。

有些图像数据库会在磁盘上写入完整的图像和该图像的一个或多个低分辨率副本，以便需要时能快速地载入这些图像。除了图像的完整原始副本外，这可能包括：用于浏览目的的很小版本，具有适合于系统打印机或网络浏览器的分辨率和色域的版本，以及其他版本。由于图像的大多数辅助版本在存储需求上要远小于原始版本，因此保留较低分辨率的副本并不会显著增加所需的总空间，而它可以大大加快访问图像的过程。柯达的 Photo-CD 是保留每幅图像的多个分辨率版本的存储格式的一个例子。在某些情况下，实际上使用一种有损压缩技术来保存较低分辨率版本的图像，以进一步降低它们的存储需求或便于网上传输。

在数据库中使用缩略图或图像的某种尺寸缩小表示的原因是，允许用户在屏幕上同时显示许多图像进而“浏览”这些图像通常是查找特定图像的基本策略。当前维护图像数据库的一种方法支持来自任何网络浏览器访问，这样，经由因特网或内联网服务器的许多浏览者便可同时查看缩略图，并具有受限的关键词搜索能力。

上面提到的关于关键词和多个数据字段的困难表明，它们几乎很少在布尔搜索中使用，进而找到一幅特定的或唯一的图像。充其量，它们可能会隔离出所存储图像中的一小部分，然后将其呈现给用户进行视觉选择。即使像自动指纹识别系统（AFIS）这样成功的和广泛报道的搜索系统，也依靠熟练的操作人员核对系统基于（每个指纹上的细节）测量值所发现的 10～20 个最好的匹配来执行最后的识别。从理想且干净的指纹中可自动提取出细节的位置和类型，但在许多情形下需要人工核对。搜索和

识别通常需要匹配 12～16 个细节。

类似地，当前的面部识别软件采用了一组 20～30 个尺寸比率，详见第 12 章。例如，从鼻子到上唇的距离与从眼睛到下巴的距离的比率，对于图像大小和面部视角相对不敏感。这些比率被用来从数据库中选择少量最佳匹配，然后将所存储的图像呈现给人工操作员。到目前为止，这些方法非常有限地成功率表明了这一任务的困难性。

浏览图像与对大多数其他类型的数据使用的搜索策略非常不同。多数计算机系统都集成有智能搜索功能（要么作为系统软件本身的一部分，要么作为基本的实用程序，要么作为应用程序的一部分），它们不仅能基于数据字段的值（如文件的创建日期）或关键词来定位文件，而且通过检查文件的内容来定位文件。例如，用于撰写本章内容的文字处理程序可搜索硬盘上的所有文件，以找到包含有词汇“图像”和“数据库”的文档，然后按名称及上下文中显示的这些词汇来显示这些文件。

这种方法可行的原因在于这些文件是以文本方式存储的。尽管可能出现特殊格式的字符（事实上，同样可对特殊格式的字符进行搜索：“查找其中包含有以不小于 14 磅 Times Roman 斜体表示的单词‘large’的文档”），但大部分文件的内容是组成单词和短语的字母的简单 ASCII 表示。文本搜索策略得到了人们的广泛使用，包括忽略大小写或要求精确匹配，以及允许使用“通配符”符（例如，“寻找包含短语‘fluorescen# dye’”的文档找到 fluorescent 和 fluorescence）。人们已为开发定位匹配的有效搜索算法进行了不少努力。

文本搜索的一个较对较新的进展是，使用自然语言规则来解释文本。这种搜索技术通过区分名词、动词和形容词来从文本中提取一些含义。这种方法，通过结合使用同义词和反义词词典来取代单词，允许通过输入一些主题句来指定目标文本，然后用搜索引擎查找并对匹配结果排序。有些程序还可分析在文档中使用的词汇，包括单词和语法类型的组合，以识别文档的可能作者。

如何将这样的技术应用于图像呢？图像数据通常作为一个像素值序列存储在磁盘上。对于许多单色图像，存储要求是每个像素一个字节，从 0 到 255 的值表示数据的灰度。对于具有更大动态范围的图像，每像素可能需要两个字节；一些计算机先存储高位字节，然后存储低位字节，而有些计算机则则与此相反。对于彩色图像，每像素需要 3 个字节或 6 个字节。对每个像素，可以用全部 3 个数值以固定的顺序存储，或者对红色、绿色和蓝色值（或某些其他色彩空间表示）单独存储整行（或甚至整幅图像）。

图像有几十种不同的存储格式。很少有图像数据库管理程序支持所有这些格式，大多数用户会根据具体需要选择一种格式，或使用某些特殊程序所要求的格式。当图像必须从某一“外来”格式导入时，通常需要使用专用程序转换它们。当不同的计算机和程序从不同源头获取图像，然后将这些图像传输给单一数据库供日后分析时，尤其需要这样做。转换过程中通常会丢失元数据。

有几个相对“标准”的格式，如在几种不同计算机平台上使用的 TIFF（标记图像文件格式），而另一些格式可能是某种特定计算机的唯一格式（如 Macintosh 上的 PICT，PC 上的 BMP）、专属于某个特定的程序（如 Adobe Photoshop 使用的 PSD 格式），或设计用于特定的应用领域（如用于医学成像的 DICOM 格式）。每家相机制造商都创造了自己独特的“原始”格式。广泛使用的 Photoshop 程序凭借自身力量已使得自己的格式成为一类标准。有些“标准”（TIFF 是一个很好的例子）有很多针对不同程序的不同选项。因此，由一个程序所写的 TIFF 文件可能无法正确地被另一个有着不同选项子集的程序所读取。

有些存储格式包括描述格式、颜色表的各类头信息，以及或多或少与文本文档格式等效的其他重要数据。有些格式会用许多方式来原始数据，其中一些方式是“无损”的，因此可完全重建图像，而有些方式是有损的，适用于降低存储需求的场合。

例如，Macintosh PICT 格式是一种无损压缩方式，它通过仅列出值一次，然后列出该值的重复次数，来表示一行有着相同值的像素。对于计算机图形学、动画和制图程序的渲染画图来说，这种“游程编码”（RLE）方法是非常有效的。它也可用于通过电话线来发送传真。但对典型的真实世界图像而

言，它并不能提供很高的压缩率，因为像素组通常不一致。

如果出现的是压缩的数据，那么计算机可能需要在对它做任何操作之前，读取整个图像文件并在内存中重建图像。但是，即使数据可以在磁盘文件内直接扫描，如何才能根据内容搜索来定位一幅特定的图像呢？一般没有沿着目标图像（大多数图像以光栅格式存储，形成行的集合）的一行的特定序列的像素值。甚至通常没有像素值的一种特定的二维模式。图像中的特征通常更不规则，并且可能会以意想不到的位置、大小和方向出现。

图像的统计平均（如图像的亮度直方图、支配色等）有时可能是有用的，但在搜索数据库中的一幅特殊图像时，通常不会计算它们。相反，如果认为这些参数很重要，那么就应在图像存储时，预先求出它们，并写入元数据中的专门数值字段中，进而能使用标准的逻辑测试来搜索它们。

一种数据匹配方法有时会应用这种类型的搜索。目标图像与数据库中的每幅图像的互相关操作，是找到与目标图像相类似的图像的一种方法。互相关的使用将在第 6 章中说明，因为它经常使用傅里叶变换来实现。这一过程可以使用专用的硬件来加速，但即便如此，搜索一个大型数据库来寻找最佳匹配仍然可能是耗时的。

该方法的一个应用是匹配监控照片来识别可能感兴趣的军事目标。例如，机场的航拍照片与飞机图像数据库互相关，每种类型以不同的方向旋转，将匹配飞机的类型及其在图像中的位置。这种技术也被用于从航拍照片中对在北极筑巢的雪雁进行计数。对更多不同的图像，或其中特征在形状、对比度或颜色方面有更大变化的图像，这种方法是不太合适的。

当图像存储在跨越几个物理盘的数据库中时，尤其是在用到可移动介质时，搜索进行时，很多图像可能无法访问到或不在线。这一问题通常的解决办法是，在主程序的一个文件中保存描述性的关键词和其他搜索字段，以及图像的缩略图。这样就可快速检索可由用户视觉检查的图像子集，为进一步的详细分析做好准备。由于该程序在其文件中备有每幅图像的位置，因此它之后可以要求用户插入特定的磁盘来加载图像。即使使用这种方法，当数据库中包含有成百上千幅图像时，搜索文件可能会变得相当大。存储搜索文件本身需要大容量的（快速）磁盘。

另一方面，搜索字段、关键词、缩略图和其他辅助信息应随图像存储而不是存储到一个中心数据文件中，还有几个原因。将这些信息存储到一个文件中，会使得该文件中包含有图像，因此可以进行整体复制。也可保存谁及何时访问了图像的记录，或保存图像自变化以来的所有记录。这些信息在重建处理图像的历史时是非常重要的。

图像数据库程序也可能需要限制对图像的访问，比如使用密码。网络文件服务器可能允许一组用户从文件中读取图像，另一组用户将图像添加到数据库中，而第三组用户处理或修改图像。由于图像很大，将它们从某个中心位置通过局域网移动到许多工作站，并不简单，尤其是在医学成像和遥感应用方面，这些应用中大量的图像将被适量的用户访问。像通过互联网的这种广域访问，会是更大的挑战。

最后，任何图像搜索程序的问题只是它被使用的方式。按照一些条件找到一幅或多幅图像通常并不是最终结果，而仅仅只是开始。图像现在如何才能被载入到程序中进行处理或测量呢？某些数据库管理程序可以充当被任何打开一个文件的程序所使用的过滤器。对加载图像而言，这很方便，但有可能不能将图像加到数据库中。其他管理程序可以定位图片并将它们复制到本地磁盘中（且可能转换它们的格式，或对其解压缩），因此用户可以更容易地将它们打开到所需要的应用中。

3.12 无损编码

前面已提到过图像压缩，图像压缩的目的是降低所存储或传输的大图像文件尺寸，是一个活跃的研究领域。本节并不回顾文献中描述的所有压缩技术，因为其中的大多数并未集成在专用的图像处理

软件包中。多数方法可分为有限的几类，具有代表性的例子如下所示。

判断图像压缩方法的标准有两个。其一是完成压缩与解压缩所需的时间及可实现的压缩程度。当图像为“实时”传输而压缩时（如在视频会议中），或通过互联网传送大图像时，这一点尤为重要。其二是图像内容保存的程度。这里主要考虑后一标准。

压缩方法之间的根本区别在于它们是否无损技术还是有损技术。无损方法是一种精确重建所有像素值的方法，而有损方法却不是这样。通常被称为图像编码而非压缩的无损方法，已存在了一段时间，最初开发这种方法的目的是用于传输空间探测器的图像。这些低功率发射机所提供的通信带宽不容许从远程相机发送许多图像，除非使用一些方法来降低每像素的位数。

早期的一种简单方法只发送每个像素与其前一像素间的差值（有时也称为“Δ 压缩”）。由于图像的大部分区域变化很小，因此降低了平均数值，进而不要求每像素 8 位，使用更少的位即可。这是图像像素高度相关的另一种说法。相邻像素间差值的直方图中，有一个靠近零的峰，而几乎没有大值，如图 3.22 所示，该图给出了不同像素亮度值的图像及其直方图。

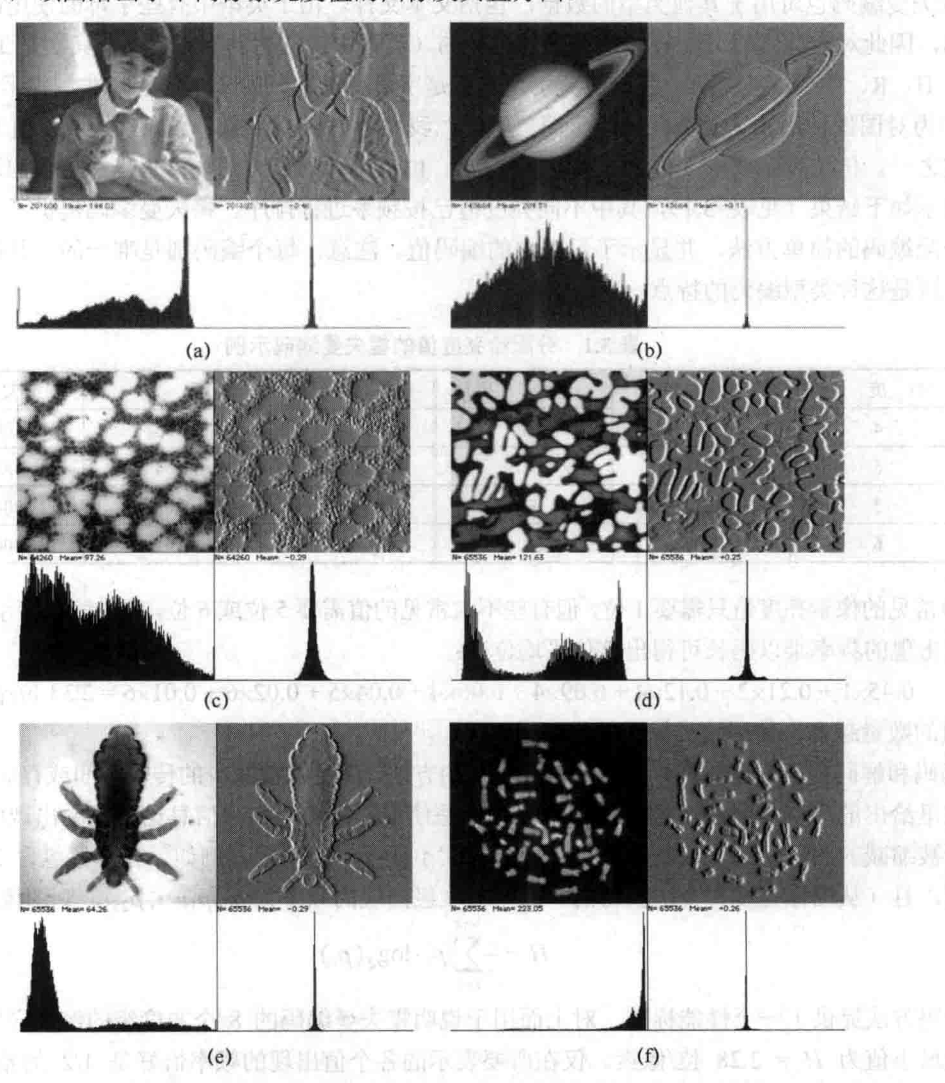


图 3.22　示例图像及其直方图，以及通过计算每个像素与其左侧邻图之间的差值的压缩结果。原图有完全不一样的直方图，但压缩图像大多数差值接近零。压缩图像以一种灰色偏移显示，以确保能见到正差值和负差值，并扩展对比度来显示像素差异。熵值列于表 3.2 中

进一步的压缩方法使用了算法，这些算法会检查前几个像素，并基于某类拟合算法预测下一个值，然后存储所得的差值。这些方法中的高级方法同样也使用前几行的像素，以进一步改善预测值，然后约简差值（Daut et al.，1993）。多数这类压缩算法最初设计用于实时图像源，并基于其时的一行像素进行某些计算。若图像存储在内存中，则可使用全部像素阵列。

另一种方法可更有效地对一个像素与其预测值间或与其前一个值间的差值进行编码。例如，在记录Landsat图像的每个波段时，对多数差值使用了一个4位数，16个可能的数字中的两个值保留为标志，以表明后面跟着一个或正的或负的更大差值。存储这些图像平均需要大约4.3位/像素，而非全都是8位/像素。

若进一步优化，用变长编码来表示差值，则对于 Landsat 数据，存储需求可进一步降低到平均大约 3.5 位/像素。广泛使用的变长编码方案之一是霍夫曼编码。在这种编码方案中，使用不同值在图像中出现的频率来给每个值分配编码。较短的编码用于频繁出现的数值，而较长的编码用于较少出现的数值。对于原始像素值和像素差值，都可以进行这种编码。

霍夫曼编码也可用于其他类型的数据，包括文本文件。由于英语中有些字母的使用此率要高于其他字母，因此对有些字母分配较短的编码是可能的（英语中最常用的字母是E，其次是T、A、O、I、N、S、H、R、D、L、U等）。字母的莫尔斯码就是字母变长编码的一个例子（非最优示例）。

作为对图像使用霍夫曼编码的一个简单例子，我们考虑一幅图像，其像素（或差值）可以取 8 个亮度值之一。传统表示要求3位（$2^3 = 8$）每像素。由该图像的直方图，可求出每个值出现的频率，并可能显示如下结果（见表 3.1），其中不同亮度值已按频率进行排序。霍夫曼编码提供了一种按该频率表来分配编码的简单方法，并显示了示例中的编码值。注意，每个编码都是唯一的，且编码序列彼此无关，这是这种类型编码的特点。

表 3.1　分配给亮度值的霍夫曼编码示例

亮度值	频率	霍夫曼编码	亮度值	频率	霍夫曼编码
4	0.45	1	2	0.06	0001
5	0.21	01	7	0.04	00001
3	0.12	0011	1	0.02	000000
6	0.09	0010	0	0.01	000001

最常见的像素亮度值只需要1位，但有些不太常见的值需要5位或6位，超过简单表示所需的3位。每个值出现的频率乘以码长可得出整体平均位数：

$$0.45\times1 + 0.21\times2 + 0.12\times4 + 0.09\times4 + 0.06\times4 + 0.04\times5 + 0.02\times6 + 0.01\times6 = 2.33\ \text{位/像素}$$

值的数量越多，使用这类编码的优势越明显。

编码和解码需要一些时间，尤其是对更复杂的方法，但足以被减少的传输时间或存储需求所弥补。这对这里给出的所有编码和压缩方法都成立，也是使用它们的理由。信息论根据所出现的值的分布，对可能被缩减的图像（或任何其他数据文件）的大小设了一个下限。例如，若图像包含 256 个可能的灰度级，且（从图像亮度直方图获得的）这些灰度级出现的频率为 $p_0, p_1, \cdots, p_{255}$，则图像的熵为

$$H = -\sum_{i=0}^{255} p_i \cdot \log_2(p_i) \tag{3.1}$$

这为编码方法提供了一个性能标准。对上面用于说明霍夫曼编码的 8 个灰度级的例子应用该式，可算出理论最小值为 H = 2.28 位/像素。仅在所要表示的各个值出现的频率恰好是 1/2 的整数次幂（1/4, 1/8, …）时，霍夫曼编码才不满足这一最优极限。但该例表明，它确实提供了一种具有适度计算需求的有用压缩度。其他可供使用的编码技术可能接近这个理论极限，但像霍夫曼编码这样的简单方法通常会好到足以被人们广泛使用。

表 3.2 列出了图 3.22 中所示几种典型图像的熵值，图 3.22 中还显示了这些图像的原始亮度直方图和水平相邻像素间差值的直方图。有些图像在原始直方图中具有较大峰值，这表明图中具有许多有着类似亮度值的像素。这会影响到熵值和最优霍夫曼码的生成。注意，随着图像外观变得更为简单且出现更大的均匀亮度区域，信息密度（*H*）会降低。如表所示，对这些图像的不同编码计算 *H* 值，表明了较大的降低，平均压缩率为 1/2。

表 3.2　图 3.22 中所示典型灰度图像的熵值（位/像素）

图	*H*（原图）	*H*（差值编码后）	图	*H*（原图）	*H*（差值编码后）
3.21a 女孩	7.538	3.059	3.21d 树状突	7.415	4.262
3.21b 土星	4.114	2.19	3.21e 小虫	6.929	3.151
3.21c 骨髓	7.780	4.690	3.21f 染色体	5.836	2.968

通过以相同的方式对邻域内所有成对的两个（或多个）灰度值的联合概率求和，可以将熵的定义扩展到覆盖邻域对。这适用于一种编码方案用于成对的或更大组合的像素，理论上这可以进一步压缩图像。

虽然对出现的所有可能灰度值对应用编码方法并不寻常，但对图像可使用一种相关的编码方案，就好像图像是文本文件那样。这种方法对文件扫描一次，以找到任何重复模式的字节值。在文本文件中，英文（或任何其他人类语言）中出现的字母模式远不是随机的。有些字母序列，如 ing、tion、the 等经常出现，且可以由单个字符替换。在特殊情形下，常用词甚至整个短语会多次出现，以允许这种表示。

使用由大小和频率选择的片段来获得最大的压缩，可以对每个文档构建此类术语的字典。字典本身和编码后的文档必须被存储或传输，但对于大文档，这种组合所要求的空间远低于原来所要求的空间。使用这种方法通常可将典型的文本文档压缩到其原始大小的一半。对于小文档，使用一种基于语言的其他样本的标准字典可能比对每个文件构建一个单独的字典更为经济。一种称为 Lempel-Ziv 技术（或它的一个变体，比如 Lempel-Ziv-Welch，LZW）的基于字典的方法，通常用来压缩文本以减少文件大小。

对图像应用同样的压缩算法，就好像它们表示 ASCII 字符那样对字节序列进行处理并不成功（Storer，1992）。多数图像中几乎没有完全重复的图案；即使少量的随机噪声也会引起像素值的变化。这些噪声波动可能并不重要，且实际上很容易认为这类“噪声”出现在照相机、电子设备和数字化过程中，而不是场景本身的一部分。但对图像编码的目标是保留所有内容，以便图像可以完全重建，且没有先验知识而丢弃小的波动。判断出像素值中的一些变化表示噪声而其他一些变化表示信息是很难的。由于这些波动，无论其来源，几乎发现不了重复的模式，且压缩后的图像通常并不明显地小于原始图像。

与这些编码方法极其相似的一种方法是游程编码，该方法已用于上面提及的一些标准文件格式中。游程编码（RLE）在所有行中寻找完全重复的像素亮度值，并用该亮度值和这种亮度值的数量来代替它。对于灰度图像，这样的行不经常出现，因此压缩率并不高。然而，对于计算机图形学、图纸和动画图像，该方法可以实现非常高的压缩率。

游程编码特别适合于二值（黑白）图像。传真机使用 RLE 经由电话线发送包含文本或图像的页面（见第 7 章中的图 7.47）。阈值处理通常用将图像简化为特征和背景的黑白表示。这样的二值图像可以由游程编码有效地表示，此外，可直接在编码后的图像进行一些度量。

台式计算机和工作站中最常使用的图像格式是图像的无损表示。有些格式会以某种规则顺序来记录所有的像素值，或许包括以像素数给出图像大小的头信息、由不同值表示的彩色查找表，或其他类

似的信息。有些格式使用适度的编码方案（如游程编码）来加速读取和写入过程并降低文件的大小。有些格式（如 TIFF 或 HDF）是可能格式的集合，该集合可能会也可能不会包含编码和压缩，并在头信息内指明特定文件中格式的细节。

所幸程序用户很少需要知道所使用格式的细节。图像存储后，要么是被同一程序以后可以读取的格式，要么是几个程序共享的标准格式，要么是各个程序通过转换后可以读取的格式。针对不同的计算机平台，人们已开发了几个这样的转换程序来读取不同格式的图像文件。遗憾的是，仍有少数公司使用非标准的专用格式。

3.13 简化的调色板

有些存储格式会设法减少每幅图像所表示的颜色数量。彩色图像占据大存储空间的原因之一是，需要为每个像素存储三色分量（通常为 RGB）。在有些情形下，每个分量会占用 1 字节的存储空间，以便为每个分量提供 256 个可能的值。而在有些情形下，会约简到 32 个灰度级（$2^5 = 32$），以将存储需求减少到 2 字节，但表面颜色的平滑变化间的幅值改变，会导致让人分心的视觉假象。在另一些情况下，每个颜色分量使用超过 8 位，有时多达 12 位（$2^{12} = 4096$），这通常是捕获彩色幻灯胶片的所有动态范围所需要的，当然这会导致存储需求的进一步增加。由于很容易对值进行存取，因此对于高动态范围图像最常见的安排是，对红色、绿色和蓝色分别使用 2 字节，或每个像素 6 字节。多数数码相机的原始格式使用这一位深。

有些图像存储时可以减少存储需求，具体取决于它们的终极应用。这使得每个像素有一个从 0 到 255 的存储值，并占用 1 字节，而存储的值表示一个有着 256 个值的表中的一项，该表则用于显示该图像。该方法看起来可能会大幅减少图像中的信息量，但它符合于一些计算机显示器工作的方式。替代使用对 R、G 和 B 采用 8 位（256 个亮度值）的 24 位显示器，一些低成本显示器使用了这样一种系统，在该系统中，256 个可能像素值中的每一个被用于选择 1600 万（2^{24}）种颜色中的一种。对应于每个像素值的调色板与图像一起存储（它仅占用 $3 \times 256 = 768$ 字节），并写入控制输出到显示器的信号的接口。因为每个像素值用于“查找”相应的显示颜色，因此这称为查找表（LUT）或颜色查找表（CLUT）。对同一 LUT 操作所产生的伪彩色显示已在第 1 章中介绍过。

若正确选择 LUT 或颜色值的调色板，则可为某些目视产生可接受的显示质量。由于选择哪 256 种颜色用于显示整幅图像（通常包含数百万像素）是该方法成功的关键，因此人们设计了各种算法来满足这一需求。（第 1 章中给出的）三色组可以视为三维色彩空间中的一个矢量，它可基于 RGB、HSI 等坐标。为图像选择最好调色板的过程包括：检查该空间中表示图像中所有像素的颜色的点，然后以一致的方法找到这些点簇，以将空间分隔为多个区域。然后使用调色板中的单个颜色来表示类似颜色点的每个集合。该过程通常称为矢量量化；搜索最优结果的迭代算法有多个，有些算法还考虑了颜色的空间分布（Heckbert, 1982; Braudaway, 1987; Gentile et al., 1990; Freisleben & Schrader, 1997; Puzicha et al., 1998; Montagne et al., 2006）。

视觉质量和速度的改进是通过使用广播视频色彩空间（YCC、YIQ 或 YUV）得到的，其中 Y（照度或亮度）轴比两个色度信号（黄-蓝色和红-绿色）更精确，因为人类视觉对亮度的变化比对颜色的变化更敏感。减小带宽在电视中传输色彩信息的道理与此相同。数值的缩放是非线性的（Balasubramanian et al.，1994），对应于 CRT 对信号强度变化的响应。换言之，调色板压缩的目标是再现图像，使得图像在电视屏幕上的视觉观察不会显示令人讨厌的人为假象（如色带）。这些方法可很好地实现这一目的。但是，若要进一步地定量使用图像，则缺失真实色彩信息就会存在问题。由于打印和视频显示的可用色域与颜色亮度响应不同，甚至可能会导致图像打印困难。

3.14 JPEG 压缩

在能容忍精确像素值的某些损失时，可考虑对图像实现更高的压缩率。讨论不同方法优缺点的文献很多（Rabbani & Jones，1991；Rabbani，1992；Kou，1995；Bhaskaran & Konstantinides，1997；Hoggar，2006；Abhayaratne，2006；Xiang et al.，2006；Ives et al.，2008）。例如，对邻域像素间的差值进行编码可以是有损的，且通过设置一个上限值可获得更大的压缩。由于多数差值很小，1 可完全表示小的差异，但只允许±7 级亮度的最大变化。这一限制将每像素位数从 8 降到 4，而无须任何其他编码技巧。较大的差异可扩展到几个像素上。当然，这可能会使重要的边缘或边界失真。

常见并在许多计算机机程序中实现的典型方法有三种。流行的 JPEG（联合图像专家组）标准广泛应用于数码相机和基于网络的图像分发。小波变换较新，但已成为新 JPEG 2000 标准的一部分（Christopoulos et al.，2000；Wallace，2001），它可使 JPEG 图像中的可视假象减至最少。首先，它为压缩使用更大的块——可选择，但通常为 1024×1024 像素，而不是原始 JPEG 方法中使用的 8×8 像素块，这种块通常会生成可见的边界（Zhai et al.，2008）。分形压缩也有许多拥护者，并声称能够通过插入超过原始分辨率限制的"真实"细节来放大图像。下面介绍每种方法并给出示例。

JPEG 技术是许多基于变换的压缩方法的典型代表。它使用一个类似于第 6 章中所示傅里叶变换的离散余弦变换（DCT）。JPEG 标准是 CCITT（国际电报电话咨询委员会）和 ISO（国际标准组织，发布号为 ISO/IEC IS 10918-1 T.81）共同努力的结果，它包括各种并不明确用于计算机存储的方法；该算法处理图像传输中可能遇到的字节流。

JPEG 变换包括以下几个步骤：

（1）使用 YUV 变换把图像分成亮度和颜色通道（如第 1 章中所示），并细分为 8×8 的像素块。如果图像在宽度或高度上不是精确的 8 个像素的倍数，则会临时填补到该尺寸。

（2）用离散余弦变换处理每个 8×8 像素块。这与我们更熟悉的傅里叶变换密切相关，唯一的不同是在离散余弦变换中，所有的值是实数而不是复数。变换为频率分量产生另一个 8×8 的值块。尽管原像素值是 1 字节 =8 位（0, 1,···, 255），但转换后的数据临时以 12 位存储，即 11 位精度和 1 个符号位。除了因这一有限表示可能出现的舍入误差外，算法的 DCT 部分不引入任何数据丢失（即原始图像可通过一个逆 DCT 由变换精确地重建）。但该步使得所需的存储增加了，而没有减少。

（3）通过除以一个对高频项提供最差精度的固定值表，将每块的 64 个系数量化到较低的精度。在大多数实现中调整"质量"因子会增大这些因子，并将更多的项降低到较低精度，或完全消除它们。这是压缩中的"有损"步骤。多数情形下，与色彩数据相比，亮度或照度的数据精度更高。这是因为对于目视图像进行压缩时，与色彩相比，我们需要图像亮度有更大的保真度。

（4）每块 64 个系数中的第一个系数是平均亮度或"直流"项。它表示的是与图像中前一块的相同项的差。块以光栅扫描整个图像的顺序列出。

（5）每块其余的 63 个系数以之字形对角线顺序扫描，从最低频率开始，并渐进到最高频率。如上所述，整个数据流用霍夫曼编码进一步压缩。这一步是无损的。

解压或图像重建过程与上述步骤相反，产生类似于原始图像的一幅图像。DCT 压缩和解压是对称的（具有相同的计算复杂性和时间要求）。一些其他压缩方法，如图像的分形压缩和电影的 MPEG 压缩，是不对称的，执行压缩所需时间比播放时解压缩的时间长得多。

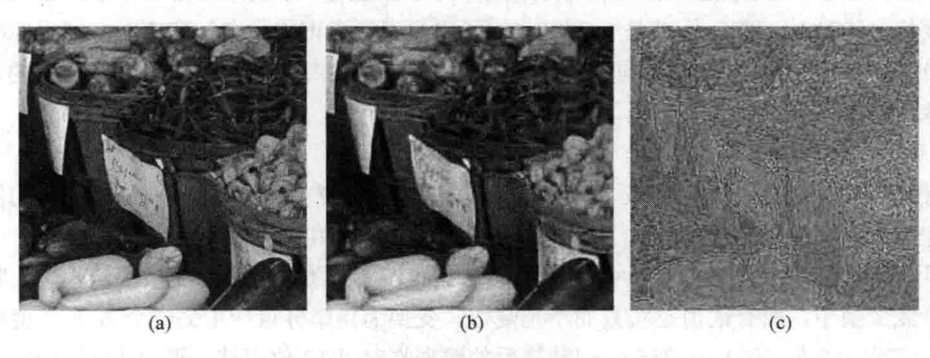
高频项的损失会导致图像出现失真和变形。由于精度损失取决于值的幅度，因此原始图像中不同 8×8 像素块的结果是不同的，且差异的确切性质会因位置不同而不同。通常，清晰的边界、边缘、角点和直线需要最高的频率来准确地再现，且正是这些显示出最大的退化。结果取决于相对于 8×8 块边

界的线或角点所在的确切位置。一个有均匀灰度值的 8×8 像素块被压缩到无损编码的单个系数，剩余的系数都等于零。来自这个均匀灰度的小偏差可能会也可能不会被保留。

JPEG 或基于变换的其他类似方法，可以使用几种方式来改善。第一，在开始之前选择表示图像的最好色彩空间。第二，替代将图像分割成非重叠的子块，在水平和垂直两个方向使用重叠块的系统，可以抑制块边界处出现的一些人为假象（Young & Kingsbury，1993）。第三，也许是最重要的，各项的量化可以更加灵活。对每个颜色通道、不同的颜色、不同的频率、不同的方向，可使用不同的缩放因子，具体取决于图像的预期用途（用于观看、印刷等）。这些方法可以在给定的压缩级别提高重建图像的质量，而重建时间不变。JPEG 2000 标准增加了几个扩展：包含了每通道超过 8 比特的图像，瓦片大小可变，而不是固定的 8×8 像素，且包含了小波变换，以作为 DCT 的一种替代选择。

JPEG 压缩或任何“有损”压缩技术应只用于目测和打印图像，而不应用于测量和分析图像。甚至对仅导致中等压缩的相对高“质量”设置，这同样适用。甚至在较低质量、较高压缩设置情形下，由于边缘和线的混叠、分辨率的降低以及对比度的抑制，图像的目视检查也可能会受到影响。

图 3.23 显示了用 JPEG 压缩并重建的一幅图像的放大片断，图像的压缩率为 10:1。主要特征仍然可以辨认，但仔细检查则会发现许多人为假象。除了块状的外观外，细节已丢失或改变。字迹不再是清晰可辨的。颜色也出现了偏移。对于分析目的，图像已严重牺牲了保真度。



(a) (b) (c)

图 3.23 一张数码照片的放大细节

所有常用的压缩方法都利用了人类视觉容忍（并察觉）的空间分辨率在颜色信息中比在亮度中要少这样一个事实。因此，这提供了一种有用的方法来检测是否应用了压缩（打开一幅压缩图像并以一种无损格式保存，当然不会恢复丢失的信息）。图 3.24 显示了图像经 JPEG 压缩后的色调通道（饱和度设置为最大）；未压缩图像的灰度值会因像素不同而不同，而压缩后的图像中，块的彩色数据空间分辨率降低了。

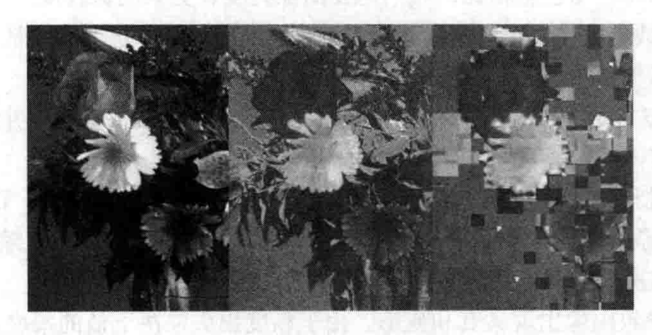


图 3.24 JPEG 压缩对颜色的影响（从左至右）：原始图像；原始图像的色调通道，其中饱和度设置成最大，以显示不同像素的颜色变化；经 30:1 的 JPEG 压缩后的色调通道，其色调值在大块上是均匀的

因为多数图像都包含有大量冗余信息，而人类视觉和理解只根据少量线索就能在熟悉的设置中识别熟悉的目标，所以JPEG压缩已基本上被接受为在消费应用和在网络上传输图像时使用。对于科学、技术和取证成像目的，由有损压缩引起的人为假象是不可接受的。图3.25显示了一个例子。胶片图像已被压缩，压缩率小于17:1，它看起来几乎无变化，但横跨胶片的平均密度图表明已完全消除了那些小峰值。

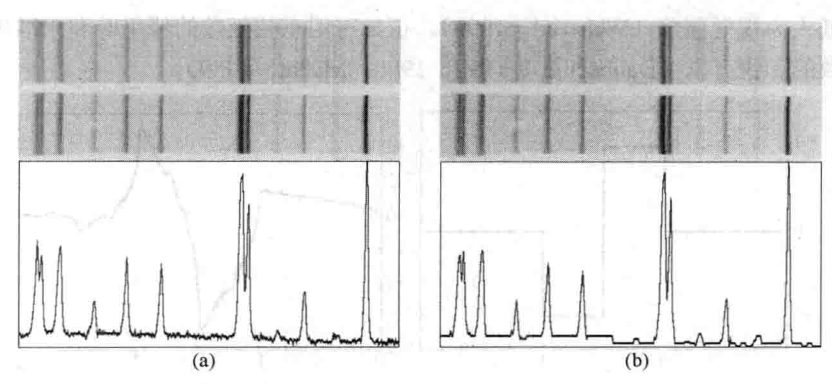


图3.25　Debye-Scherer X射线胶片用JPEG压缩前(a)和压缩后(b)的图像（从136.3KB压缩到8.1KB）。垂直线的位置和密度提供有关晶体结构的信息。横跨图像的密度分布图表明，频谱中的统计噪声已随同一些小峰一起被移除，而压缩则引入一个新的虚假峰

判断压缩图像和复原图像的质量并不容易。通常需要计算原始像素值和重建值间的统计差异，但不提供与人们判断及图像测量和分析所需非常吻合的测量。如第2章中所讨论的，人类视觉对像素间的相同绝对变化的反应不同，具体取决于它们是处在图像中的亮区还是处在图像中的暗区。由于人类视觉的非线性响应，暗区中差异通常更为明显。平滑或强烈且有规律图案区域的差异，通常认为比随机区域的差异更为重要。像素差异导致的边缘或边界移位，与特征内部的亮度和颜色变化相比，通常认为更为不利。

人们通常采用图像质量评级来比较不同的压缩方法，但即使仔细比较，结果仍依赖于照明条件、图像上下文及疲劳程度。大量人工比较的结果使得我们选择了现在广泛使用的JPEG标准。JPEG方法是众多依赖于首先将图像从熟悉的空间域变换到按照一组基函数分离所呈现信息的另一空间的方法之一。

3.15　小波压缩

将图像变换到另一个空间的优点是，可以一种不同的方式组织信息。傅里叶变换（和JPEG压缩中使用的类似DCT）按照频率和方向将呈现的信息进行分离。若使用这些函数的一个完整集合，则可以准确并完整地重建原始图像，因此该变换本身并不是一种有损方法。但在使用函数的完整集合时，需要说明幅度和相位的数量正好占用与原始图像一样多的存储空间。事实上，因为幅度和相位值是实数，而原始像素值是需要较少存储空间的整数，所以它们通常要占用更多的存储空间。

除了用于傅里叶变换或DCT方法的正弦函数外，还有许多其他基函数集。有些基函数提供更有效的图像表示，其中的多数项很小，而有些基函数更易于计算。最近较流行的方法之一是小波变换，它提供了一些计算上的优势，甚至可以用光学镜头和掩模光学地得到（McAulay et al.，1993），或用滤波器组以电子方式得到（Drygajlo，1992；Strang，1996）。

小波变换在不同尺度提供了图像的渐进或多尺度编码，这比傅里叶变换这样的加窗方法更为灵活。小波由一组归一化正交函数组成，图像被投射到这些正交函数上（Chui，1992）。这些函数是局部的，

而不像正弦函数那样无限延伸到图像之外，所以小波变换往往能更好地处理区域与图像的边缘，并且将它用于压缩避免了在 JPEG 压缩中经常看到的“块效应”或“接缝现象”，在 JPEG 压缩中，进行离散余弦变换之前图像被划分成 8×8 的像素块。

常用几种不同的小波函数。最简单的是 Harr 函数，它只是一个可以跨越图像移动并水平地展开的方阶，如图 3.26 所示。其他小波函数，如所示的 4 阶 Daubechies 小波，虽不太明显，但以同样的方式运行且可以降低人为视觉假象（Welstead，1999）。正如可用正弦函数的傅里叶和来准确重建任意函数一样，小波函数的和也可以（Daubechies，1992，1996；Mallat，1989）。

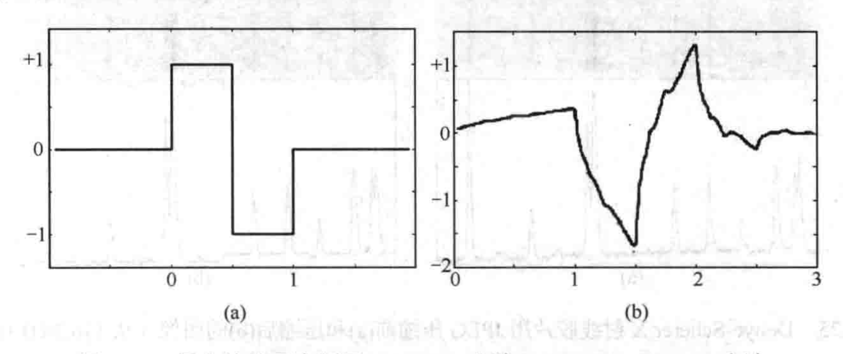


图 3.26　母小波的两个例子：(a)Haar 小波；(b)Daubechies D4 小波

为了解使用 Haar 函数的小波变换过程，我们考虑下面的过程。首先，找到整个图像的平均亮度并将该值放到变换左上角的单个像素中（完成时将与原图一样大）。现在把图像再分成 4 个象限并找到每个象限的均值与总体平均值之间的差。将这些值置入一个 2×2 像素块。将两个左侧象限中的每一个象限与右侧对应象限之间的差值放至一个像素中，并将类似的垂直差值放到一个像素中，如图 3.27 所示。通过细分象限重复这个过程，并继续下去直到达到单个像素时为止。

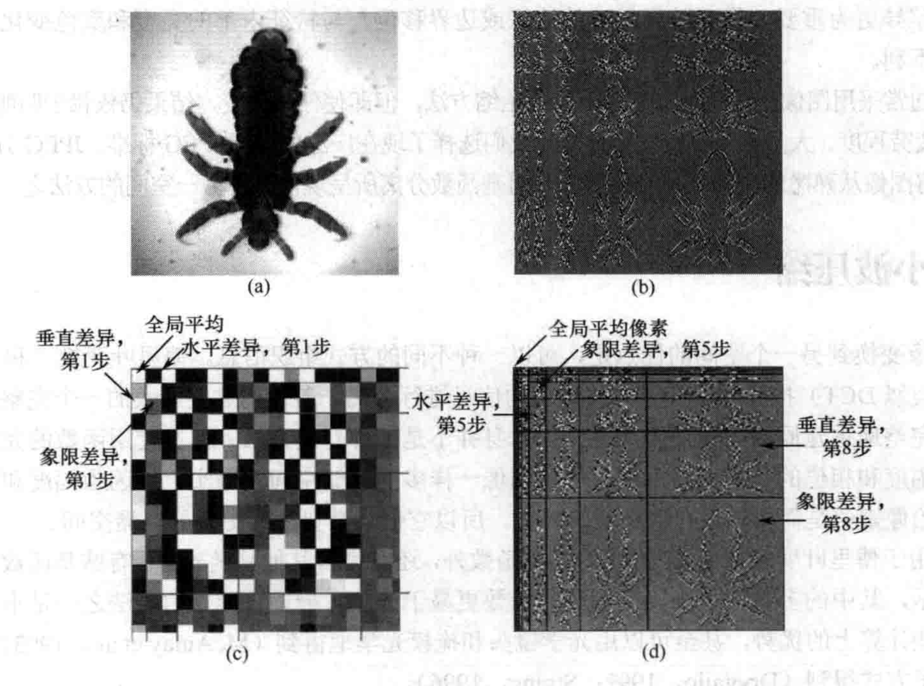


图 3.27　应用到一幅示例图像的小波变换过程。变换(b)由不同尺度下平均值之间的差值组成。它们标记在变换(c)和整个变换(d)的左上角

显然，在变换中含有较高频率差异的正方形和矩形（它们位于右下方）比那些与低频信息相对应的数据更大并且表示了更多的数据。细分过程相当于在原始图像中用越来越高的频率（较短的距离）进行工作，且可以用选定的频率来重建原始图像。在图 3.28 中显示了重建图 3.27 中图像的过程，它通过加回在逐渐更小的阶的像素之间的差异来实现。该过程精确、完全无损地重建了原始图像。该方法对很多时间相关和空间相关信号的应用利用了其有效及简洁的优势。在某些情况下，系数的分析可以提供有关图像内容的信息，就像对傅里叶变换常做的那样。

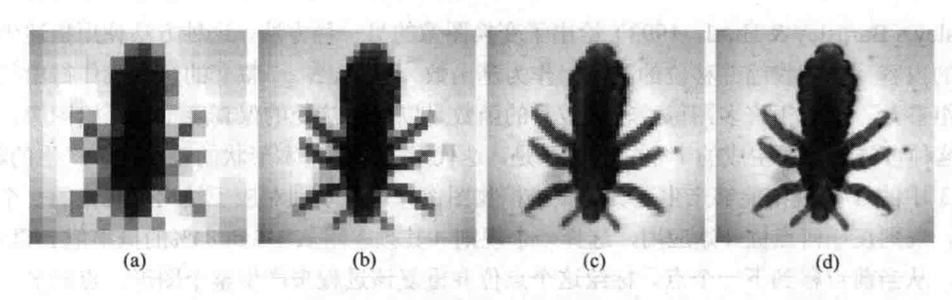


图 3.28　重建（逆小波变换）逐渐加入了每个更高频率项。在本例中，显示了第 4 步、第 5 步、第 6 步和第 7 步。第 8 步恢复了原图像

原理上，所有这些变换都是无损的，且所占空间与原始图像一样多（但和傅里叶变换不同，Haar 小波变换由整数组成，因此不会占用更多的空间）。为实现压缩，变换中的系数被量化，就像对 JPEG 余弦变换那样。图像的有损压缩是通过除法使得变换中的不同项的幅度变小来实现的，所以它们可以用更少的位数来表示，并删去了那些非常小的项。这可以降低所需的空间，但也在重建的图像的质量上产生了损失，并引入了不同的人为假象。

用小波变换对所选定的频率进行滤波是通过从该变换的相应部分中减少或消除一些值来实现的。这类似于傅里叶空间的滤波，将在第 6 章中讨论。

图 3.29 显示了有损小波压缩的效果。虽然在重建的图像中人为假象在视觉上并不那么令人讨厌，仔细观察原始图像与压缩图像之间的差异，会发现许多细节上的不同，如在边缘位置的颜色变化和位移，这严重影响了定量测量。

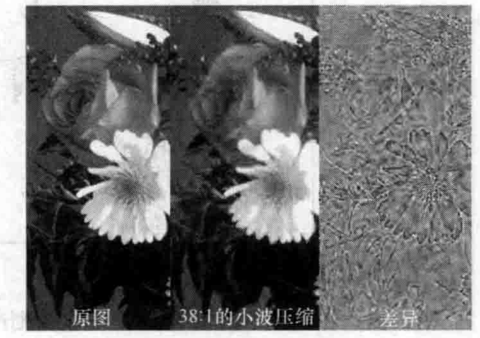


图 3.29　由小波压缩引起的损失：从重建图像中减去原始图像而显示出差异，图像中具有更多细节的区域的差异表现最为明显

对图像进行压缩的理由是，来自变换特定项的损失对于人类观看者来说，可能并不可见或至少并不感到讨厌，因为每一项所表示的信息是遍布整个图像的，且对任何特定的特征可能没有贡献。但并非总是这种情况。选择小振幅的项往往意味着从图像中删除高频信息，这可能对定义边缘和边界尤其重要。正如在 JPEG 余弦变换中视觉上最反感的缺陷是“块状”的人为假象，因此小波变换最讨厌的方面是在图像的“繁忙”部分更多的细节被抑制，所以很难预测什么细节会保留而哪个将不复存在。

使用有损方法的压缩往往适合于在特定应用中使用的图像，如在印刷中印刷设备本身会对可能重现的保真度设定一个上限。前面已提及，压缩的一个应用可能是在数据库中存储图像的多个副本，对于具体的用途，每幅图像具有适当的质量用于特定的目的，这些目的包括如打印或为了搜索和识别以缩略图观看。确定由不同程度的图像压缩产生了多少质量的损失同样是重要的，以便避免可能改变对重建图像分析的人为假象。

常用压缩的另一个应用（此时为小波压缩）是指纹建档，由执法机构用于鉴定。指尖上的纹路间距在个体之间仅有少量变化，所以保留了相应的频率，而丢弃了更低的频率（如图像中的阴影）和更高的频率（如噪声或灰尘）。

3.16 分形压缩

Barnsley（Barnsley & Hurd，1993）给出了变换图像的另一种方法，这种方法使用描述为图像的自仿射变形的内容（原件收缩并移位的副本）作为基函数。原理上，一整套的这些操作得出与原图像一样大的一组参数。但对于许多图像，较少数量的函数足以用可接受的保真度重建原始图像，提供有效的压缩。这种自仿射图像生成的一个知名例子是，迭代组合相同基本形状的较小副本产生的蕨类植物。

表 3.3 中的 4 个规则能够产生逼真的蕨类植物图像。每个规则对应于结构子元素的一个旋转、位移和收缩。规则在任何点位开始应用，选择一个规则（其频率用从 1%到 84%的概率值 p 显示），然后根据规则，从当前点移到下一个点。标绘这个点位并重复该过程来产生整个图形。点越多，结果的定义越好，如图 3.30 所示。具有 20000 个点的整个蕨类植物显示了整体形状的自相似性。为检查目的而放大部分图像时，就会需要更多的点。最后，放大率由计算中所用值的数值精度设置，如图所示。

表 3.3　蕨类植物图像的变换（基函数）（见图 3.30）

规则	变换	规则	变换
1 ($p=0.840$)	$x'=+0.821x+0.845y+0.088$	3 ($p=0.075$)	$x'=+0.076x+0.204y+0.494$
	$y'=+0.030x-0.028y-0.176$		$y'=-0.257x+0.312y-0.133$
2 ($p=0.075$)	$x'=-0.024x+0.074y+0.470$	4 ($p=0.010$)	$x'=+0.000x+0.000y+0.496$
	$y'=-0.323x-0.356y-0.260$		$y'=+0.000x+0.172y-0.091$

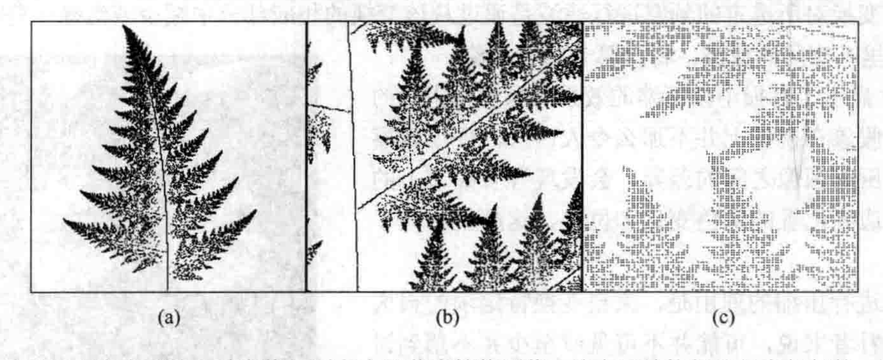


图 3.30　使用正文中介绍的 4 种变换规则产生的蕨类植物图像。放大后结构仍然是自相似的，只是由于计算的有限数值精度限制，其舍入了 100 倍放大图像中的值：(a)具有 20 000 个点的蕨类植物图像；(b)放大 5 倍的图像；(c)放大 100 倍的图像

从数学的观点来看，这 4 个变换是可以按比例（p 或概率值）相加来产生整个目标的基函数或映射。同样的原理已应用到带有灰度或彩色像素的摄影图像的压缩中，此时它称被为拼贴定理（Collage Theorem）。因此，必须要有这样的基函数或规则用于自仿射变形，且原理上它们是可以找到的（Barnsley et al.，1986；Barnsley，1988；Khadivi，1990；Barnsley & Sloan，1992；Barnsley & Hurd，1993）。然而，找到正确映射并不容易。知道这样的函数必然存在对找到它们并未提供什么线索。每个映射由平移、缩放、旋转和扭曲组成。找到它们的一种专有方法已被授予专利（Barnsley & Sloan，美国专利号 5065447）。进行分形压缩有正式的流程（Fisher et al.，1992；Wu et al.，2005），虽然结果不一定是最优的。业已证明，也可以使用非线性自仿射变换，且可能更加有效。

这种技术被描述为“分形压缩”，因为重建是迭代执行的（如蕨类植物所示），且提供了更多的细节。事实上，该方法可以继续下去，以产生细节尺度上比发现基函数的原图像更精细的重建图像。这些细节看起来非常令人印象深刻，因为放大图像从不显示表明分辨率有损的平坦或平滑区域。当然，细节并不真实，因为这是在假定大尺度图像中呈现的任何模式，也会在所有更精细尺度下以逐步减少振幅的方式呈现这一条件下产生的。

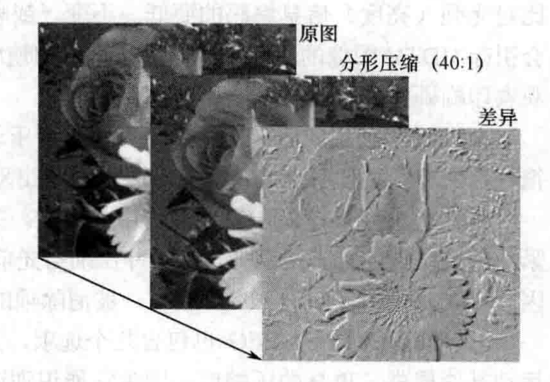


图 3.31　由分形压缩引起的损失，包括颜色的改变、特征位移，以及插入了原图中不存在的细节

与 JPEG 和小波方法不同，分形压缩的是非对称的。对图像压缩所需的时间通常远远大于重建它所需的时间。分形压缩也有特有的人为假象。图 3.31 显示了一幅压缩率为 40:1 的分形压缩图像。从重建图像中减去原图像表明色彩已经明显改变，特征发生了位移，且插入了原图中并不存在的细节。

3.17　数字电影

使用个人计算机来生成、编辑、在网络上发布并显示数字电影已变得很常见。两个比较突出的标准（或标准集）是苹果公司的 Quicktime 和微软的 Windows 视频。这两种方法都允许使用不同的压缩技术，通常称为 CODEC（COmpressor-DECompressor）。CODEC 是前述压缩算法之一的程序实现（JPEG 是由现有 CODEC 提供的方法之一）。大多数 CODEC 只在软件中运行，而少数 CODEC 使用额外的硬件来获得实时的速度。对于 JPEG，这既可以是一个执行离散余弦变换的数字信号处理器芯片，也可以是一个特别为 JPEG 算法设计的专用芯片。

将整部电影压缩到单张 DVD 上、将视频上传到网上、广播 HDTV 图像，或允许以较低带宽网络传输图像的数字视频会议时，JPEG 方法并不能为数字电影提供太大的压缩，因为 JPEG 方法并不考虑序列图像相当大的冗余。多数情况下，电影胶片或录像带上的一系列图像，随着时间的推移，较大区域的像素（噪声除外）要么完全不会改变，要么以相对平滑和连续的方式改变。压缩这些内容是可能的，且可以容忍高压缩比，因为对于这些区域，重建图像所需的质量并不高。人类视觉往往会忽略变化不大或无变化的图像部分。

MPEG（运动图像专家组）标准遵循和 JPEG 同样的方法。这种方法的目的是，将图像压缩到足以允许图像在电视台使用的带宽内向 “用户”传送视频和高清电视（HDTV），或将整部电影存入 DVD 中。当然，对于更多的技术目的，如跟踪图像序列中的特征，或测量密度或颜色随时间的变化，不同需求可能会对可接受的重建质量设置不同的限制。

对于图像序列，另一种压缩基于连续帧之间的相似性，而 MPEG 增加了几个额外的步骤来减少必须传输的数据量。这种方法寻找连续图像中的相似像素块，即使它们稍有移动。通常每秒平均只有 2 帧是整体发送的。其余帧均编码为与之前帧的差。这种方法可以在保证“可接受”视觉外观前提下，让整体压缩比接近 200:1。

典型的高质量 HDTV 图像以每秒 30 幅图像的速率呈现 1920×1080 像素，总的数据率超过千兆比特每秒。图像采用逐行扫描（而不是隔行扫描）、方形像素及 16:9（而非用于 NTSC 制式视频的 4:3 屏幕高宽比。典型广播电视台的带宽非常小。重要的是要记住，这些压缩方法的标准是，当重建图像

显示在电视屏幕上时，图像在视觉能被人们接受。例如，MPEG 标准对色度（颜色）信息编码的精度比对光强（亮度）信息编码的要低。不变（或略有变化的）块的更新不太频繁，从而在传输有困难时会引起 HDTV 图像的块状外观。为技术目的使用图像时，或以其他形式展示图像时，如通过反射光来观察印刷硬拷贝，可能就不那么宽容了。

MPEG 中使用的音频压缩自身可用于音乐录制（流行的 MP3 格式）。对保真度的要求是高频必须很好地保留，以便听众能从弱音小号中分辨出双簧管。

绝大部分运动图像压缩方法使用关键帧，关键帧使用与静止画面相同的方法压缩。然后对序列中紧随着关键帧的每一幅图像，确定并压缩与先前图像中的差异。对于这些差异图像，值的幅度减少了。因此，在执行图像的进一步压缩时，被消除项的数量会显著增加并达到更高水平的压缩。

和 JPEG 那样，MPEG 也包含几个选项，其中一些需要更多的计算，但提供更大的压缩。例如，运动补偿提供了更高的压缩度，因为它能识别连续图像中的整体平移（如相机缓慢地在一幅场景上移动时），并在将一幅图像中的位置与前一帧进行比较之前对此进行调整。MPEG 方法是非对称的，即压缩原始数据比解压它来重建图像需要更多的计算。

这种压缩方式的结果之一是，其目的在时间上一直是向前的。由关键帧及后续的差异，可以重建每一幅接下来的图像。但是不容易重建之前的图像，除非返回到最近的一个之前的关键帧并从那里向前操作。从原理上来说，在影片中每个场景开始时的一个关键帧应该足够了。实际上，关键帧会被周期性地插入。

序列图像的其他压缩方法正在开发之中。一种方法是预测向量量化，这种方法试图定位每幅图像中的边界，跟踪这些边界的运动，并采用预测来生成后续图像，以便使需要编码的差异最小。报道称 8 位灰度图像序列可以压缩到平均数据速率小于 0.5 比特每像素（Nicoulin et al., 1993；Wu & Gersho，1993；Wen & Lu，1993；Hwang et al.，1993）。分形压缩也已推广到处理图像序列（Li et al.，1993）。

运动图像的高压缩比适合于视频会议，此时的图像质量只需显示谁在讲话及他们正在讨论什么。对于许多消费应用，最终图像仅在中等分辨率的屏幕上观看（如手机或上传到网络的影片片段），因此高压缩率可得到足够好的图像质量。对电视观众的测试表明，最重要的是声音的质量，而个别图像中的明显缺陷并不那么令人讨厌。从网上下载的许多高压缩率的电影包含极端的人为视觉假象，但观看它们的那些人却认为是可以接受的。

然而，对于大多数技术和取证应用，静态图像压缩产生的各类人为假象是不能接受的，且由时间压缩结果引起的额外情况会使事情变得更糟。打算对来自一个序列的图像进行分析的用户，应从完全没有压缩的图像开始，且试验表明压缩图像对于特定目的可以接受时，才采用这些特定的压缩方法。

第 4 章 修正成像缺陷

本章讨论最先一组图像处理运算，即用来校正图像获取时，由于传感器不完美、光学系统的局限性、不足或不均匀光照，或不理想视角所导致缺陷的那些程序。需要强调的是，这些修正是在图像已被数字化和存储后应用的，因此无法传递通过优化或校正最初获取过程来实现的最高质量结果。修正通常应在增强前或测量后应用。

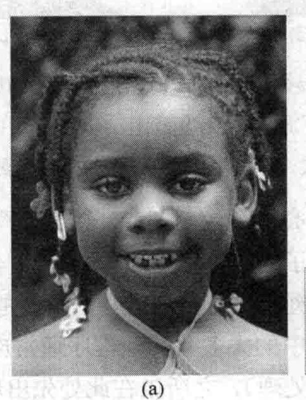
当然，获得最佳质量的图像有时是不切实际的。若相机在实时或场景变化前仅能收集少量光子，则图像中出现的噪声不能通过获取和增加更多的光子或视频帧来消除，而需要其他的降噪手段。若不能将光源控制成正好在中心且垂直于表面（如太阳），或若表面是曲面而非平面，则图像可能具有非均匀光照，这必须在之后修正。若视角不能调整（如空间探测器或卫星轨道），或表面不规则（如金属断裂中的情况），则场景中的有些部分是透视收缩的，在比较大小或测量距离时，必须要考虑到这一点。

即使是在典型的实验室设定中，使像显微镜这样的仪器保持在理想对齐状态下可能也会非常耗时，且仪器保持稳定状态并用很长时间去收集暗淡的图像可能是非常困难的。因此，为了方便和快速，需要在理想情况下可实现的图像质量和使用图像处理方法来完成这些修正之间进行折中。在获取第一幅航天探测器图片时，人们认可了这种类型的修正，而当时这种修正在中速计算机上需要较长的处理时间。现在，我们在台式计算机或笔记本上几毫秒时间就可以完成这样的修正，因此可以实际地应用到日常的成像需求中。

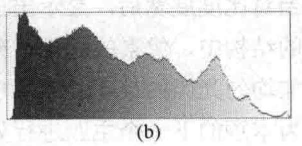
4.1 对比度扩展

第 1 章解释了在图像的典型数字化过程中，会产生从 0（黑色）到 255（白色）的值，产生 1 字节（8 位）值，或对彩色图像的红、绿、蓝分量都产生 1 字节。若相机和数字化仪具有更高的精度，这些值可能会有 10、12 甚至更大的精度位，且通常每个占用 2 字节。当这是相机传感器输出的全动态范围时，实际的图像数据就必须覆盖全部范围。在许多情况下，所记录图像具有更小的亮度值范围，原因在于这些亮度值位于中间（中间灰度值），或接近亮端或暗端。

图像直方图（记录每个可能亮度级的像素数量的图形）是用来观察图像对比度的有用工具。图 4.1 给出了一个例子，其中直方图覆盖了整个动态范围，且表现出了良好的对比度。没有像素完全是黑色的或白色的。多数廉价相机在检测器中的模数转换器的协助下，使用比存储结果图像（通常为 8 位）更多的精度位（如 10 位或 12 位），都能很好地捕捉日常生活场景中曝光良好的图



(a)



(b)

图 4.1 亮度值覆盖整个范围而未对黑色或白色进行限制，所以是一个良好曝光调节的例子：(a)一个完全曝光的图像；(b)及其亮度直方图

像。因此，我们可以调整最终值以覆盖整个亮度范围。对于保存原始图像数据的专业数码相机，原始文件随后的转换可以实现相同的结果。

图像亮度偏差的固有范围比相机、后续电子设备和数字转换器的动态范围小很多时，实际的数字范围要比 0～255 这一范围小得多。图 4.2(a)给出了一个例子，其中标本是组织的切片，显示在亮场显微镜下。显微镜的光照和切片的轻污染产生了非常小的全对比度。窄峰及直方图两端的空白区域表明，许多可能的亮度级未被用到。

通过拉伸对比度，将像素值被重新分配到覆盖整个范围，可以改进现有结构的能见度。图 4.2(b)显示了这一结果。映射是线性且一一对应的。这意味着原始图像中最暗的像素被分配为黑色，最亮的像素被分配成白色，而中间灰度值则被赋予了新值，这些新值是黑与白之间的线性插值。原始图像中具有某个灰度值的所有像素，在结果图像中被分配了相同的新灰度值。

与图像一起画出的直方图现在显示了展开到整个可用亮度范围的用于灰度级的像素计数。但许多灰度值仍然在直方图中显示出零计数，表明没有像素取那些值。灰度值的重新分配增强了所呈现像素的视觉对比度，但并未增强辨别未记录在原始图像中的灰度的细微变化能力。同时，它也放大了与原始图像的亮度差异和噪声。

图 4.2(c)显示了用相机和数字转换器的整个范围得到的相同视野。这可能需要调整光照、相机的增益或曝光时间等，并可能需要在设置上反复试验。如果可以显示一个实时直方图，那么这一过程就可以简化。不同结构的平均亮度与图 4.2(b)所示的类似。然而，所有 256 个可能的灰度值在图像中都存在，而样本密度的细微变化现在可以分辨出来或在样品中测量。

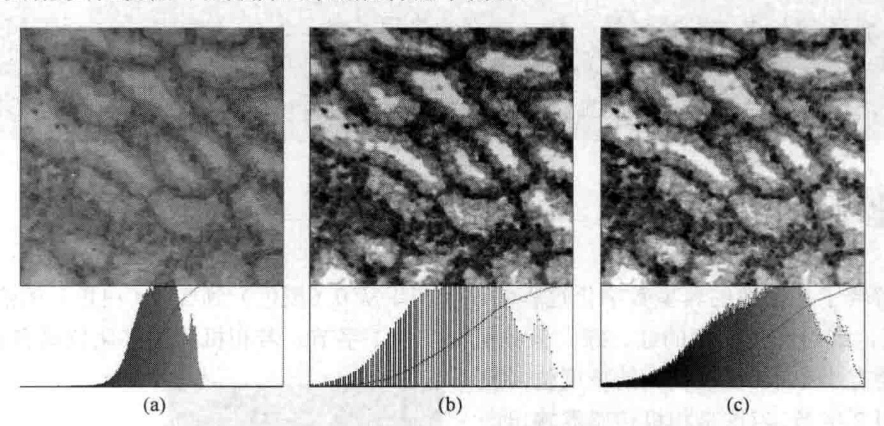


图 4.2　染色组织光学显微镜图像：(a)对比度很低的原始图像及其亮度直方图；(b)通过控制显示屏的亮度范围，线性扩展了对比度范围，但直方图中出现了空白；(c)使用最佳光照和相机曝光所拍摄图像的视觉效果，注意直方图中无空白。红色线条显示了累加或积分直方图，它从 0%上升到 100%

这一问题并不限于明亮的图像。图 4.3(a)显示了源自扫描电子显微镜的一幅较暗图像及其直方图。当对比度展宽后，显示出了集成电路的结构［见图 4.3(b)］，但这也增大了噪声的能见度。在这个相同的结构中，像素的随机“斑点”变化在亮度上应该是理想均匀的。尽管从工作流程的观点来看，降噪、非均匀光照的校正和本章中的大多数其他程序通常应该在对比度调整前开始进行（图像噪声的问题作为本章的下一个主题进行处理），之所以在此处先出现，是因为它使得对其他校正的需要更加明显，且在任何情况下这是大多数用户首先进行的调整。

所示例子是一种极端的情形，但调整光照、相机增益等来准确填入可用的像素深度（可以数字化或存储的灰度级的数量）往往是不实际的。此外，亮度范围增加太多，可能会引起在范围暗端或亮端

的像素值超过数字化和存储容量且被限制到极限值，这也会导致信息丢失。图 4.4 显示了一个夜景的例子，其中建筑物表面的亮度范围较好，但亮光比相机的白色极限更亮，因此被限制到最大值 255，而建筑物周围的暗色区域曝光不足并被限制到零（黑色），失去了任何可能已经存在的细节。

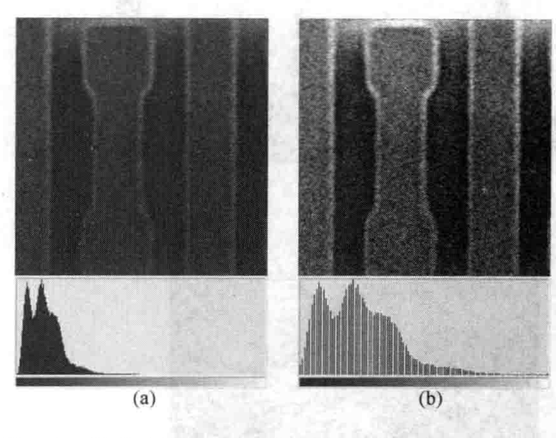


图 4.3 集成电路深色扫描电子显微镜（SEM）图像及其直方图(a)。线性对比度扩展后(b)，特征和噪声更为明显，且直方图显示出空隙



图 4.4 夜景中明亮的灯光被限制为白色而阴影被限制为黑色。由于限制而丢失的信息无法恢复

对彩色图像应用对比度扩展时，正确的步骤是将它从 RGB 格式转换到 Lab 或 HSI 彩色空间，然后在扩展亮度的同时保持颜色信息不变，如图 4.5 所示。这会阻止红色、绿色和蓝色直方图被线性扩展到整个范围所引起的颜色偏移，如图 4.6 所示。如果原始图像主要仅由几种颜色组成，或者是以一种颜色为主，那么这样的偏移特别常见。

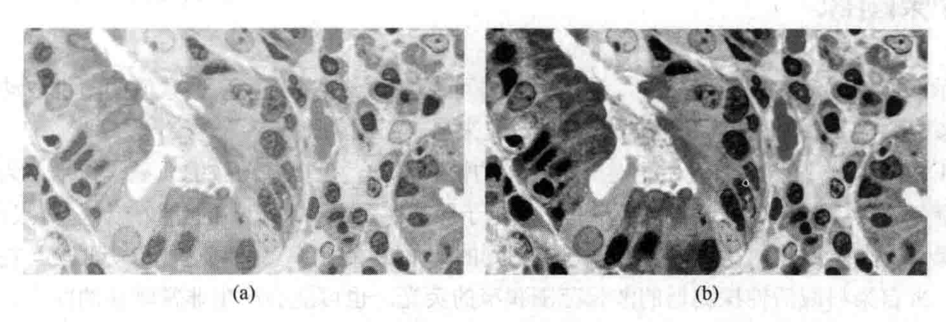


图 4.5 染色组织光学显微镜图像(a)及转换到 HSI 空间扩展对比度且只线性扩展亮度的结果(b)

如果图像具有足够的不同亮度级来揭示标本中的重要特征，那么线性对比扩展是一种以增强观看者视觉辨别能力的有用和可以接受的方法。更重要的是，通过将亮度范围轻微不同条件下获取的图像都调整到相同扩展后的对比度范围，就可使得直接对比它们成为可能。当然，这仅在结构的最亮和最暗区域存在于所有图像和视野中时才有效。

还可以对像素的亮度值执行称为点运算的其他操作，像素分配的新值只依赖于原来的像素值，而不依赖于任何相邻像素值，且是一对一的，即有着单个亮度值的像素被赋予了另一个值。但该过程可能不是线性的。例如将亮度转换成密度时，会是对数关系。对于彩色图像，可用传递函数来校正由于光源色温引起的颜色失真（如第 1 章中的图 1.22 所示），或校正卫星图像的大气散射和吸收。这种函数可以用一个数学函数或一个查找表来实现。

第 5 章举例说明了很多更为灵活的此类对比度操作，进而增大图像中结构的能见度。

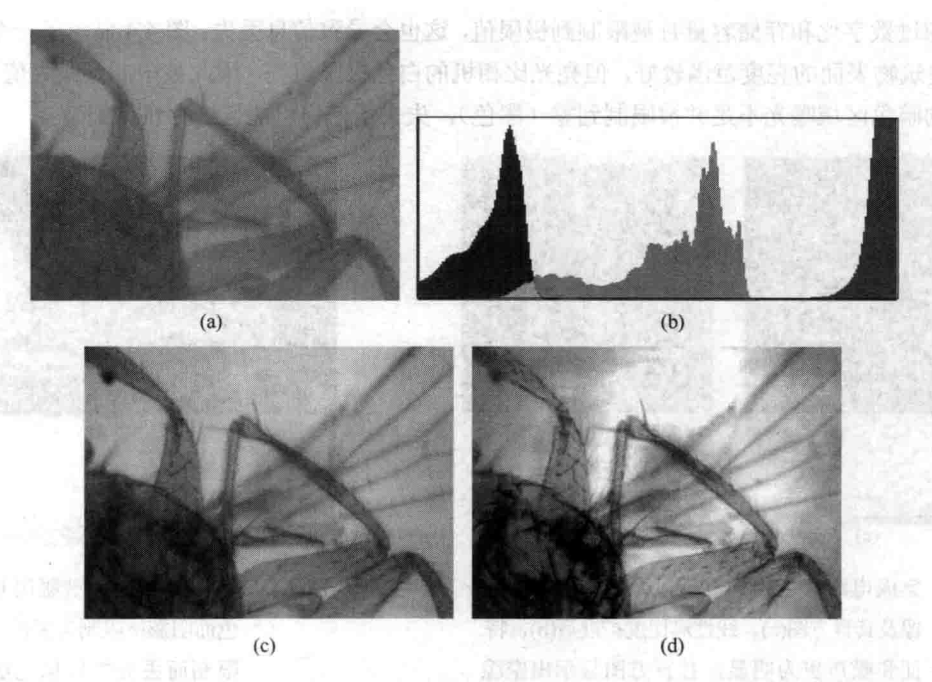


图 4.6　昆虫光学显微镜图像(a)及其红、绿和蓝直方图(b)。扩展亮度而保持色调和饱和度不变(c)，扩展了可见的对比度而未改变颜色。分别将 RGB 值扩展到全部范围，由于红、绿和蓝的比例发生了改变，因此产生了图像中的颜色偏移(d)

4.2　带噪图像

以上示例所示的对比度线性扩展，通常会使得可能的噪声的能见度（像素值的随机波动）增加。噪声是图像中的重要缺陷，其呈现形式有多种，且来自于不同的源。

在第 1 章中，图 1.27 举例说明了对一定数量的帧取平均，可提升图像的质量（技术上为信噪比）。一个不可避免的噪声源是少量入射粒子（光子、电子等）在图像检测器中的计数统计。对来自 SEM 的 X 射线图像情况尤其如此，其中入射电子与检测到的 X 射线的比率可能高达 100 万比 1。在荧光光学显微镜中，来自染料或活性探测器的波长范围狭窄的荧光，也可能会产生非常暗淡的图像，此时需要结合一系列持续时间很短的图像，来测量作为时间函数的活性。在天文学领域，于合理曝光时间内捕捉更多光子的需要，已导致人们创建了越来越大的望远镜。

在对一幅图像扫描或数字化期间，光源或检测器的不稳定也可能会导致带噪图像。由于计数统计，该噪声模式本质上可能与高斯噪声不同，但它仍然表现为场景的均匀区域中亮度的变化。一个常见的例子是，针尖场致发射电流的变化所形成的场致发射 SEM 图像中的噪声。以秒级的典型时间常数，电子发射可能会从一个原子移到另一个原子，产生百分之几的束电流变化。在所观察的图像中，将这种波动效果降至最小的常用方法是，采用比波动时间更长或更短的扫描时间。

在使用荧光照明，特别是使用用于照亮底片的灯箱获得的图像中，可以看出类似的效果，它源自相机的曝光和荧光灯管闪烁之间的拍频或干扰。这种闪烁要比白炽灯泡的闪烁大许多，因为后者的热惯性降低了光发射中交流电引起的变化。当噪声具有在时间上不随机且不出现“正态”统计行为的特点时，更难对噪声进行有意义的数字描述。但使用许多相同的技术（尽管效率较低）通常可以降低噪声。源于电气干扰或振动的周期噪声源，通常可以用图像的傅里叶变换表示来消除，详见第 6 章。

由于每个探测器灵敏度的固有变化及放大率的变化、电子转移过程中的损耗及随机热噪声，随机噪声也会源于相机芯片本身。有些噪声是横跨图像的固定模式，有些噪声则是叠加在图像数据上的随机"斑点"噪声。

在一幅图像可以表示实际中获得的最佳质量的假设下，本节探讨抑制噪声以提升所现特征的可见性并对特征进行区分的方法。所有这些方法的基本假设是，图像中像素的尺寸远小于任何重要的细节，且对出现的大部分像素而言，它们的相邻像素代表相同的结构。基于这些假设，就可以应用不同的求平均法和比较法。

与经典的图像取平均或长曝光时间中的固有假设（不同时间于每个位置的像素读数，表示视场中的同一结构）相比，这些假设几乎完全相同。这表明对像素读数求平均或在时间上对像素读数积分，可以降低随机噪声。当信号以上面提到的其他方式变化时，可以使用其他方法如中值滤波（在下面讨论）来降噪。这些方法与这里描述的空间比较类似，不同之处在于，它们使用的是在每个位置的测量的时间序列。

使用帧平均合并来自相机的许多连续读数的降噪方法，与读出连续读数前在相机内部对电荷积分来降噪的方法相比，有很大的区别。后者主要用于天文学、荧光显微术和图像非常模糊的其他应用中，且由于相机对入射光是"开放"的，因此有时也称为"凝视"模式。这两种方法看上去可能是等效的，因为两种方法都会把一定时间内的入射信号加在一起。然而，理解它们是将要考虑的两种完全不同的噪声源非常重要。

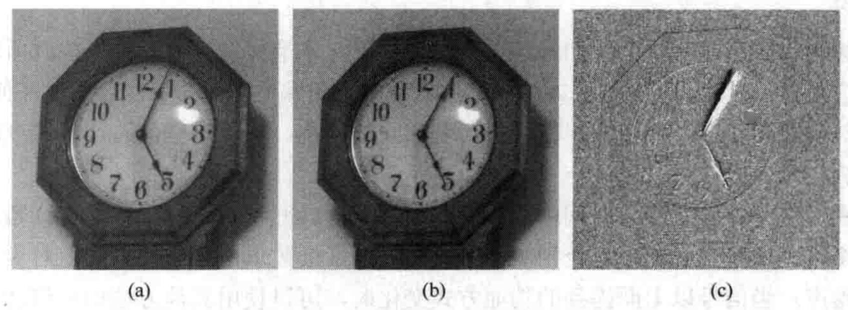
在使用"凝视"模式的相机中，于 CCD（电荷耦合器件）阵列的每个晶体管中收集的电子，包括那些由入射光子产生的电子，也包括设备本身的暗电流所产生的电子。暗电流受热噪声的强烈影响，因此热噪声会在室温下移动电子。使用珀耳帖制冷器将芯片冷却几十度，或使用液氮或液氦将芯片冷却几百度，可以显著地降低这种热噪声。与可见光相机相比，红外相机的降温幅度必须更大，因为可见光子本身的能量更低，因此信号电子的产生需要较少的能量，进而室温下会出现更大的暗电流。

用于凝视应用或长时间曝光的相机，通常会规定暗电流半填充芯片动态范围所需要的工作时间。对于使用珀耳帖制冷器冷却后的相机，有用的工作时间可能是几分钟。对于用于专业天文学并冷却到非常低温度的高质量装置，有用的工作时间可能是几小时。在凝视模式下以该长度时间收集一幅图像，会将暗灰度级提升到中间灰度，且任何真正的信号会叠加到该背景上。由于热电子的产生是一个统计过程，因此并非所有像素都有相同的背景灰度。背景的波动表示图像的一类噪声，这类噪声在非常模糊的图像中通常会处于支配地位。多数情形下，这种噪声源一定要能与读出噪声进行比较。

所有相机都有一些读出噪声。在 CCD 相机中，为能以电压形式读入计算机，来自每个晶体管的电子必须转移多次。将电子从图像的一边移动到放大器需要更多的转移，因此图像的一边产生的噪声要比另一边的大。在隔行扫描相机中，每个检测器旁会改用单独的一行晶体管，这在一定程度上降低了读出噪声。当然，来自其他相关电子设备（时钟信号、从相机到数字化仪的连线、计算机本身的电信号拾取等）的额外噪声源可能会使信号退化更多。即使通过精心设计和正确布线最小化了这些噪声源，在每次读出、放大和数字化相机图像时，信号上仍会叠加不能再降低的噪声。

该噪声通常是随机的，因此增加特定像素的亮度有可能使这一噪声降低。图 4.7 显示了使用质量较好数码相机和标准光照获得的两幅连续图像。除了指针的运动，两幅图像看起来基本相同。但从一帧图像中减去另一帧（并如前所述扩展对比度），就可看出图像中存在的像素噪声及指针的运动。对于数码相机，常用方法是使用合适的曝光时间来获取单幅图像，然后将其数字化后传送到计算机或存储

芯片中。但没有冷却设计的有些相机会获取几幅曝光时间较短的图像，然后在计算机中对这些图像进行平均。

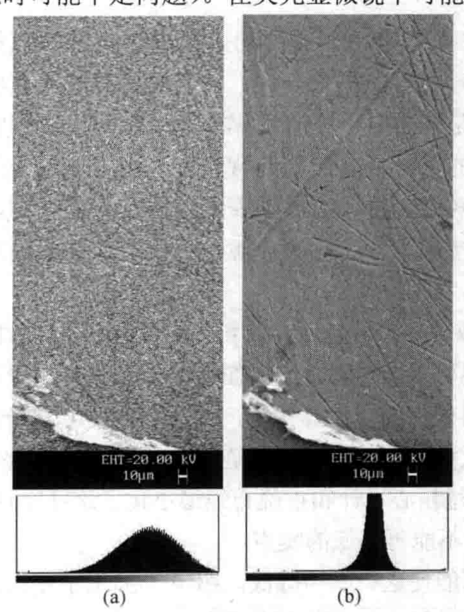


(a) (b) (c)

图 4.7 数码相机拍摄的同一视图的两幅图像，以及它们之间的差（为显示细节，放大了像素并扩展了对比度）。除了与指针运动相关的差异外，其他差异源于相机的随机噪声，且这些差异在红色、绿色和蓝色通道中不尽相同

使用摄影机将许多帧放在一起求平均，可在使随机噪声部分消失同时，继续增强信号。小芯片上晶体管的较小阱尺寸会使得长时间曝光不切实际，因此帧平均可以降低图像中噪声源的相对幅度，但不能消除它。

对于非常暗淡的图像，最佳方法是在相机芯片内对信号进行积分，但不允许任何像素达到饱和状态。然后读出数据一次，以便产生最小的读出噪声和最佳图像。对于明亮图像，或对于仅能以视频速率工作的相机，或对于其检测器具有较小阱尺寸或电子容量的相机，可能更倾向于平均足够数量的连续帧来减少由于随机噪声引起的像素变化。该积分可能会在帧抓取器或在计算机程序中完成。帧抓取器通常会受板上存储器数量的限制，但可以在数据无损的情况下收集每一帧。计算机程序更灵活，但将帧加在一起的时间需求可能会导致丢弃一些视频帧，使得整个获取需要更长的时间（图像不随时间变化时可能不是问题）。在荧光显微镜下可能会出版褪色，因此需要尽可能快地收集所有光子，方法要么是帧平均，要么是对非常暗淡的图像使用相机的“凝视”模式。

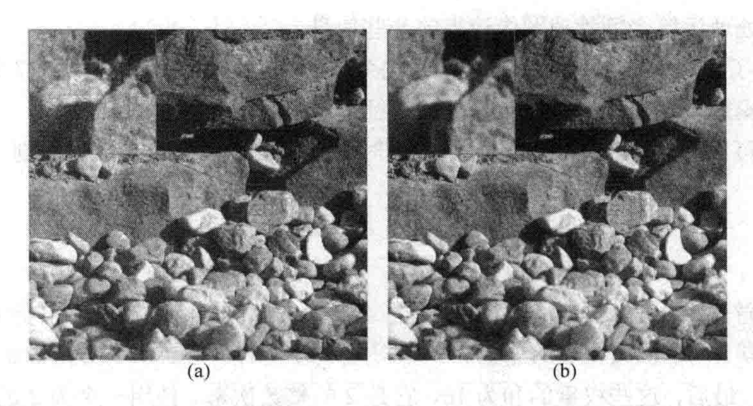


(a) (b)

图 4.8 金属表面带有刮痕的 SEM 图像：(a)1 秒扫描图像及其直方图；(b)20 秒扫描图像及其直方图

图 4.8 比较了两幅 SEM 图像，这两幅图像是以不同的扫描速率拍摄的。快速扫描图像每像素收集的电子很少，因此具有较高的随机噪声电平，进而模糊了图像的细节。将扫描速度从 1 秒降到 20 秒的结果表明，信号量得到了增长，且噪声得到了降低。直方图表明均匀区域内亮度的变化减少（峰变窄）了，这就是细节可见度得到改进的原因。

许多数码相机允许设置一个等效的 ASA 速率，该速率对应于胶片感光度。ASA 速率越高，曝光时间就越短，和/或用来获取图像的光圈就越小。这就要求更大的信号放大增益，产生更高的随机噪声电平。图 4.9 表明，使用更长的曝光时间（将相机的 ASA 设置从 1600 减小到 100）降低了噪声并提高了图像质量。图 4.9(a)所示带噪图像也是下面所示各种降噪方法的起点。



(a) (b)

图 4.9 收集更多的信号来降低噪声：(a)以 ASA 1600 设置、1/800 秒曝光得到的图像；(b)以 ASA 100 设置、1/50 秒曝光得到的图像

4.3 邻域平均法

空间平均的最简形式是，对图像内每个小区域中像素的亮度值求和，然后除以该邻域中的像素数，并使用结果值来构建一幅新图像。图 4.10(a)表明，本质上这生成了具有较少像素数的图像。块大小是 3×3，所以是 9 个像素值求和。对于该图像的随机噪声，图像质量或信噪比的改进是 9 的平方根（或因子 3）。但图像的横向分辨率受到了严重影响，且图像中的较小结构不再能被单独地识别。

实现邻域平均更为常用的方法是，用其本身及其相邻像素的平均来替换每个像素。这常被称为“核”运算，因为这种方法可以概括为区域内像素值的总和乘以一组整数权重。这一过程也称为卷积，且可在傅里叶空间中等效执行（详细讨论见第 6 章）。

$$P_{x,y}^{*}=\sum_{i,j=-m}^{+m}W_{i,j}P_{x+i,y+j}\Big/\sum_{i,j=-m}^{+m}W_{i,j} \tag{4.1}$$

式（4.1）表明计算是在一个 $2m+1$ 维的方形上进行的。这是一个奇数值，因此邻域的尺寸范围从 3×3 增大至 5×5、7×7 等。也可使用非方形区域；对于较大的邻域，近似的圆比方形更好。简单邻域平均所用的权重矩阵 $\boldsymbol{W}$ 中只包含 1 项，因此对于 3×3 区域其可以写为

$$\begin{bmatrix}1&1&1\\1&\mathbf{1}&1\\1&1&1\end{bmatrix}$$

约定是，这些系数乘以围绕中心像素的像素的值，总和除以权重之和（此例中是 9）进行归一化，并将该值写入中心像素的位置来形成一幅新图像。

实际中，因为存储空间不是无限的，所以通常要将新图像写入与原图像相同的存储空间中。但在这样做时，对和值使用原始像素值而非对邻域已计算出的新值很重要。这就需要在处理过程中复制图像的几行。

邻域运算，包括核乘法，通常会对称地应用到每个像素的周围。对于那些到图像边缘距离近于邻域一半的那些像素，这会带来新的问题。处理这一问题的方法有多种，包括对边缘或角上的点设计特殊的非对称核或规则，同时假设图像边缘是镜像对称的，以便图像内的每行像素可以复制到图像外，从图像区域内的像素值推断出边缘之外的像素值，或假设图像是环绕的，即左边缘和右边缘、顶部边缘和底部边缘都是连续的。有时会用到一种更为简单的方法，即只对图像中没有边缘冲突问题的部分进行处理，这会使得沿图像边缘宽度等于邻域半径的几行像素不被处理。这些方法都不能令人完全满

意，且通常大多数处理运算会牺牲掉图像边界的一些信息。

图 4.10(b)显示了使用 3×3 邻域平均进行平滑的效果，图 4.10(c)显示了使用 7×7 邻域进行平滑的效果。邻域越大，降噪效果越好，但伴随有明显的特征边缘模糊现象。

使用非 1 权值可以降低这种模糊现象，同时可对邻域平均过程施加更多的控制。例如，值

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

有几个优点。首先，中心乘以原始像素内容的 4 是最大的因子，它使得中心像素在平均中处于支配地位并能减少模糊。4 个正交相邻的 2 值和 4 个对角线分布的 1 值表明了对角像素（以因子 $\sqrt{2}$ ）远离邻域中心的事实。最后，这些权重的和为 16，它是 2 的整数次幂。使用一个为 2 的整数次幂的数字，可使得在计算机中通过移位运算来进行乘除操作成为可能，且这种操作速度非常快。

本书第一版出版时，由于计算机能力的限制，因此将重点放在图像处理运算的有效实现上。为加快实现的速度，同时使乘积之总和不超过计算机中单精度整数值的上限 32678（2^{15}），权重值通常设为较小的整数。2 的整数次幂得到了人们的支持，因为选择这种值的主要原因在于实现的有效性，而非理论要求。所幸的是，这些限制不再重要。现代计算机对高精度算术和浮点乘法的运算速度更快，甚至对于现在常见的百万像素图像，对多数目的而言其速度都已足够。

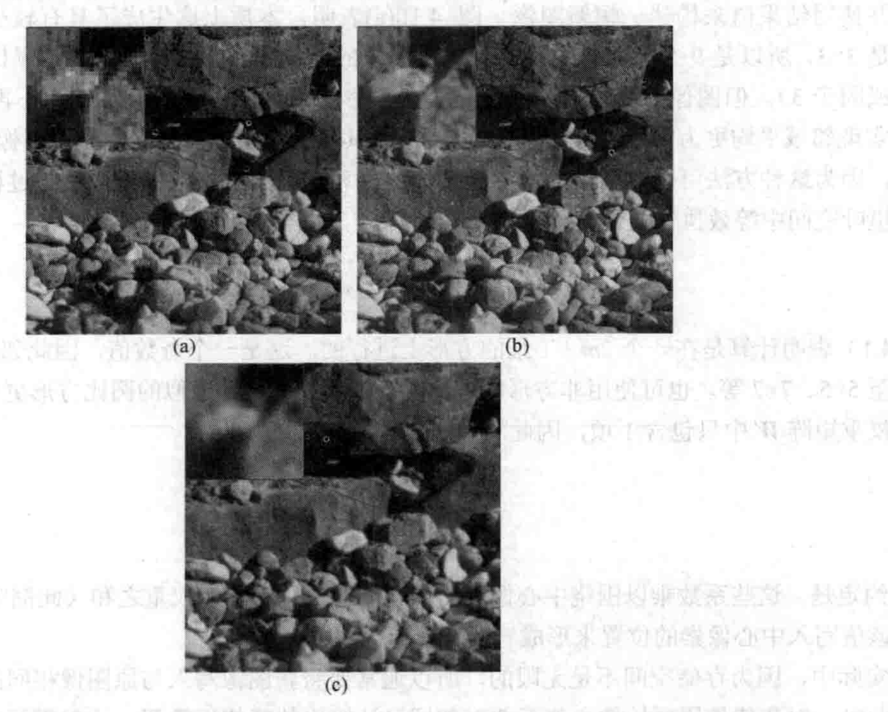


图 4.10　邻域平均［与图 4.9(a)所示原始图像比较］：(a)3×3 邻域平均；(b)每个像素被半径为 1（3×3 像素）的邻域平均代替；(c)每个像素被半径为 3（7×7 像素）的邻域平均代替

第 6 章中关于图像频率域处理的内容表明，在频率域中可以有效地分析这些核的平滑性质。权重核的一个非常有用的“形状”是高斯核，这是一组沿任何行、列或通过中心对角线逼近高斯函数曲线的权重。它由以像素尺寸表示的标准差来表征，计算公式为

$$G(x,y,\sigma)=\frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \tag{4.2}$$

式中，x 和 y 是到核的中心的距离，单位为像素。核的大小是中心像素两边的标准差的 3 倍，故添加另一行值将往矩阵中插入微不足道的小数字。这些核的标准差为包含这些系数 68%的综合幅度的半径（以像素计），在核以三维权重值图形画出时，则为面下的体积，如图 4.11 所示。这是常用标准差定义的二维推广；对于一维高斯分布，曲线下面积的 68%位于±1 个标准差内。

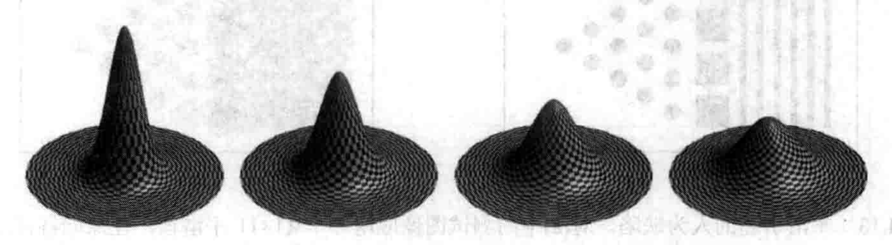


图 4.11　高斯平滑核中的权重值图。随着标准差增加，峰值下降，而权重值之和（面下体积）保持不变

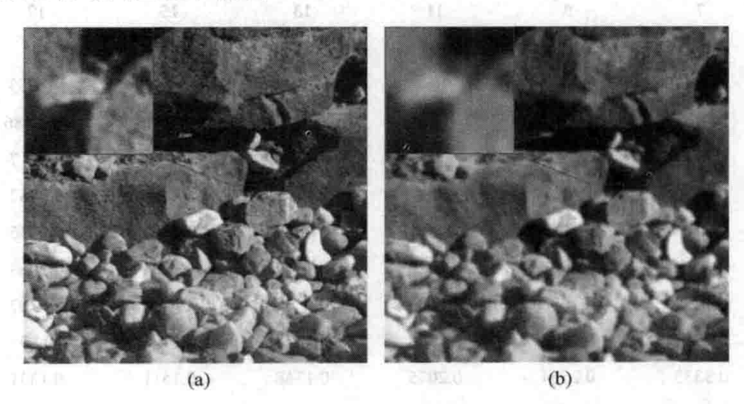
因为目标是逼近平滑的高斯分析曲线（Russ，1995d），故选择一组整数来实现高斯分布是困难的。有些系统允许将核权重输入为实数，而许多系统在用户输入标准差时，会生成所需的高斯值。

在高斯滤波器情形下，通过将该运算分为两个更为简单的运算，可以加速平滑运算。替代使用整个方形矩阵的权（对于一个 15×15 的核，这需要 225 次乘法和加法），该滤波器可以分为具有一个线性权重矩阵的垂直高斯模糊（15 次乘法和加法），以及一个水平高斯模糊（也是 15 次乘法和加法，共计 30 次乘法和加法）。式（4.3）用于计算权重的值，这些值可再次标定和舍入为整数。

$$G(x,y)=\frac{1}{\sqrt{2\pi}\cdot\sigma}\exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{4.3}$$

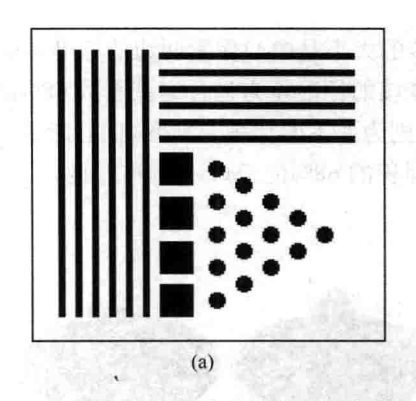
有些类型的卷积核可以用这种分离方法实现，但多数不能分离，且为理解算法的目的，最好考虑整个权重矩阵。在关于图像处理的后续章节中，给出了核的其他用法，这些用法中的权重幅度上不对称，且不全是正数。核的实现保持不变，只是在出现负权重时，通常会执行归一化，方法是仅除以正值之和，或仅除以最大的正值（因为这些情形下所有权重之和通常为零。对于图像中噪声的平滑，通常仅涉及正的权重。

图 4.12 显示了使用具有 1.0 像素的标准差（一个 7×7 的核）和具有 3.0 像素的标准差（一个 19×19 的核）的高斯核对图 4.9(a)进行平滑后的结果。这类平均确实降低了图像中的可见随机噪声，但也模糊了边缘，偏移了边界并降降低了对比度。当两个邻近的结构以在两者之间创建一种明显特征的方式平均时，甚至可能会引入人为缺陷。图 4.13 给出了这样一个例子，其中测试图案中的线条被 11×11 的平均窗口模糊了，导致它们之间出现了假线。



(a)　　(b)

图 4.12　高斯平滑的结果［与图 4.9(a)所示的原始图像比较］：(a)标准差为 1.0 像素；(b)标准差为 3.0 像素


(a)

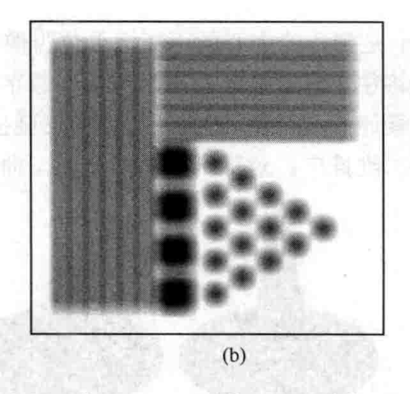

(b)

图 4.13 平滑引起的人为缺陷。对(a)中的测试图像应用一个 11×11 平滑核，在原始特征间产生了明显的线条，如(b)所示

如第 6 章所述，频率变换与一个卷积函数相乘，等同于在空间域应用一个核。最常见的噪声滤波方法是移除高频信息，这表示与随机噪声相关联的逐像素变化。对图像应用一个滤波器或掩模来降低或消除高频信息，然后再进行反变换，可以实现这种消除。所得结果与在空间域中得到的结果相同。

事实上，对于较小的核，或那些可分为垂直和水平两部分的核（如高斯核），空间域中像素值的乘法和加法的运算更快。对于比 11×11 更大的核（该限制一定程度上取决于计算机的特殊性，图像是否可以完全放入内存中或者必须从磁盘存取，是否可以使用多个处理器，等等），频率域方法更快。然而，由于这两个过程在数学上是完全相同的，因此基于空间域的权重核可能更易于理解这一过程。

在第 5 章关于图像增强的介绍中，给出了负权重值围绕一个正的中心峰值的核的用法。这一改动的目的是锐化边缘，并避免因消除噪声所引起的一些模糊。同样的方法一直用于一维信号曲线的平滑，如 X 射线衍射图案、谱或时变电子信号。这通常是使用 Savitsky and Golay（1964）的拟合程序实现的。为此目的给出的系数表用做权重系数，就像上面给出的权重系数那样，只是它们仅对一个维度进行操作。这一过程相当于将数据点最小二乘拟合到一个多项式。平滑后的曲线保持了阶的幅度，但同时消除了噪声。表 4.1 列出了用于第二次（二次项）和第四次（四次项）幂多项式的系数，以便拟合扩展到从 5 个点的邻域到 19 个点的邻域。这些曲线如图 4.14 所示。同样的方法可以扩展到二维情形（Edwards，1982）；图 4.15 显示了应用一个 7×7 的 Savitsky-Golay 二次多项式去平滑一幅噪声图像的结果。

表 4.1 Savitsky-Golay 拟合系数

二次多项式拟合							
5	**7**	**9**	**11**	**13**	**15**	**17**	**19**
0	0	0	0	0	0	0	−0.0602
0	0	0	0	0	0	−0.065	−0.0226
0	0	0	0	0	−0.0706	−0.0186	0.0106
0	0	0	0	−0.0769	−0.0118	0.0217	0.0394
0	0	0	−0.0839	0	0.038	0.0557	0.0637
0	0	−0.0909	0.021	0.0629	0.0787	0.0836	0.0836
0	−0.0952	0.0606	0.1026	0.1119	0.1104	0.1053	0.0991
−0.0857	0.1429	0.1688	0.1608	0.1469	0.133	0.1207	0.1101
0.3429	0.2857	0.2338	0.1958	0.1678	0.1466	0.13	0.1168
0.4857	0.3333	0.2554	0.2075	0.1748	0.1511	0.1331	0.119
0.3429	0.2857	0.2338	0.1958	0.1678	0.1466	0.13	0.1168

（续表）

二次多项式拟合							
5	**7**	**9**	**11**	**13**	**15**	**17**	**19**
−0.0857	0.1429	0.1688	0.1608	0.1469	0.133	0.1207	0.1101
0	−0.0952	0.0606	0.1026	0.1119	0.1104	0.1053	0.0991
0	0	−0.0909	0.021	0.0629	0.0787	0.0836	0.0836
0	0	0	−0.0839	0	0.038	0.0557	0.0637
0	0	0	0	−0.0769	−0.0118	0.0217	0.0394
0	0	0	0	0	−0.0706	−0.0186	0.0106
0	0	0	0	0	0	−0.065	−0.0226
0	0	0	0	0	0	0	−0.0602

四次多项式拟合							
5	**7**	**9**	**11**	**13**	**15**	**17**	**19**
0	0	0	0	0	0	0	0.0458
0	0	0	0	0	0	0.0464	−0.0343
0	0	0	0	0	0.0464	−0.0464	−0.0565
0	0	0	0	0.0452	−0.0619	−0.0619	−0.039
0	0	0	0.042	−0.0814	−0.0636	−0.0279	0.0024
0	0	0.035	−0.1049	−0.0658	−0.0036	0.0322	0.0545
0	0.0216	−0.1282	−0.0233	0.0452	0.0813	0.0988	0.1063
0.25	−0.1299	0.0699	0.1399	0.1604	0.1624	0.1572	0.1494
−0.5	0.3247	0.3147	0.2797	0.2468	0.2192	0.1965	0.1777
1.5	0.5671	0.4172	0.3333	0.2785	0.2395	0.2103	0.1875
−0.5	0.3247	0.3147	0.2797	0.2468	0.2192	0.1965	0.1777
0.25	−0.1299	0.0699	0.1399	0.1604	0.1624	0.1572	0.1494
0	0.0216	−0.1282	−0.0233	0.0452	0.0813	0.0988	0.1063
0	0	0.035	−0.1049	−0.0658	−0.0036	0.0322	0.0545
0	0	0	0.042	−0.0814	−0.0636	−0.0279	0.0024
0	0	0	0	0.0452	−0.0619	−0.0619	−0.039
0	0	0	0	0	0.0464	−0.0464	−0.0565
0	0	0	0	0	0	0.0464	−0.0343
0	0	0	0	0	0	0	0.0458

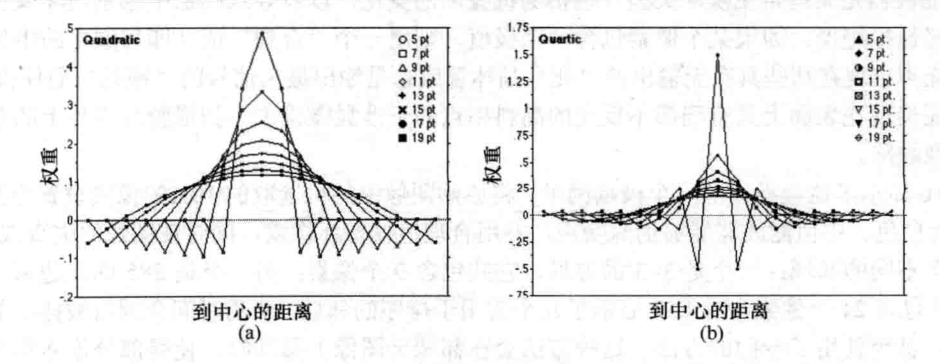


图 4.14　二次和四次多项式最小二乘拟合的 Savitsky-Golay 线性平滑权重

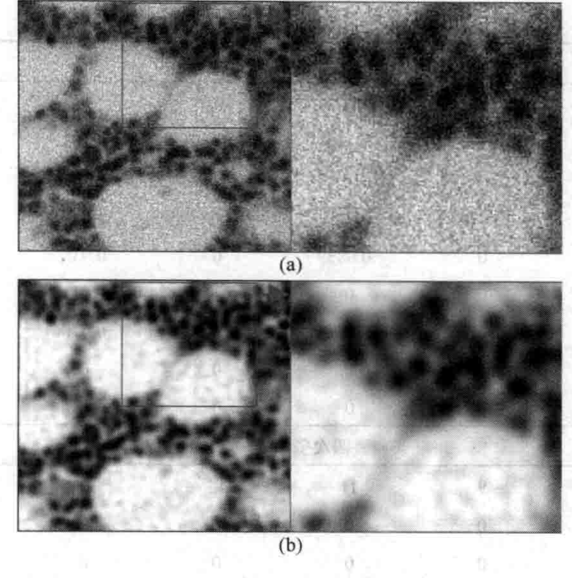


图 4.15　使用 7 像素宽的二次 Savitsky-Golay 拟合的平滑：(a)带有细节放大的原始图像；(b)带有细节放大的结果

4.4　邻域排序

平滑滤波器确实可以降低随机噪声，但其基本假设是邻域中的所有像素都表示相同值的多个样本，即它们都属于同一结构或物体。很明显，在边缘和边界处这并不成立，且上面所示的所有平滑滤波器会模糊边缘并使边缘移位，而我们并不期望这种结果。对于某一给定数量的降噪，高斯滤波器会产生最小的边缘模糊，但不能完全消除模糊。

使用加权核将邻域中的像素放在一起进行平均是一种卷积运算，如上所述，这种运算在频率域空间图像处理中有其对应运算。这是一种线性运算，该运算会使用邻域中的所有像素，且原始图像中的信息不会丢失。在空间域，对邻域也可执行其他的降噪处理操作。这些操作都不是线性的，且在频率域没有等效的操作。

这些方法中，最广泛使用的方法是根据亮度对邻域中的像素进行排序。排序后像素的中值可用做新的中心像素值。与核运算的情形类似，邻域排序用于生成一幅新图像，且对围绕每个像素的邻域进行排序时，只使用了原始的像素值。

中值滤波器是某些常见噪声类型（包括随机叠加的变化，以及导致图像中像素毁坏或丢失的冲激噪声）的极佳拒绝器。如果某个像素包含一个极值，则用一个“合理”值（即邻域中的中值）来代替它。这类噪声出现在那些具有无输出的“死”晶体管或总是输出最大信号的“锁死”晶体管的 CMOS 相机中，还出现在表面上具有局部不反光的高斜率点的干涉显微镜中。扫描胶片底片上的灰尘也会导致这种图像缺陷。

图 4.16 显示了这类噪声的一个极端例子。将原始图像中随机选取的 10%的像素设置为黑色，另外 10%设置为白色。中值滤波器能够消除噪声，并用合理值代替坏像素，同时使图像的失真或退化最小。使用了两个不同的邻域：一个是 3×3 的方形，它共包含 9 个像素；另一个是 5×5 的八边形（近似圆）区域，它共包含 21 个像素。图 4.17 显示了几个常用于排序的邻域。所需时间会因将被排序的值的个数快速上长，甚至使用了专门的方法，这种方法会在邻域于图像上移动时，使得部分像素集单独排序，只需要一些额外的像素比较（Weiss，2006）。方形邻域与近似为圆形的邻域相比，更易于实现，但在邻域尺寸增大时，使用圆形区域对于产生各向同性的结果非常重要。

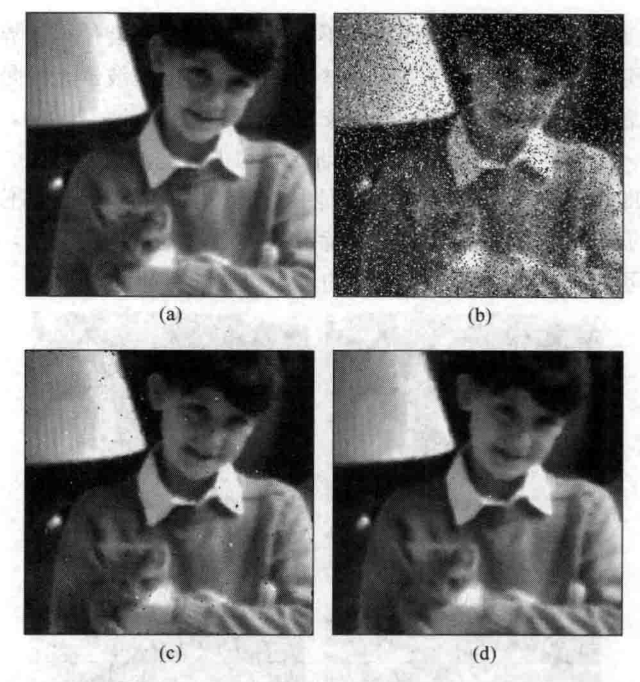


图 4.16 用中值滤波器移除散粒噪声：(a)原始图像；(b)选取图像(a)的 10%的像素并设置为黑色，同时选取另外 10%的像素并设置为白色；(c)使用 3×3 方形区域对图像(b)应用中值滤波的结果；(d)使用 5×5 的八边形区域对图像(b)应用中值滤波的结果

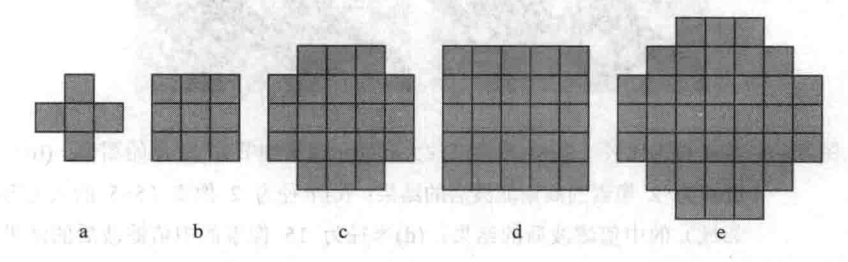


图 4.17 用于中值滤波的邻域图形：(a)4 最近邻交叉；(b)包含 9 个像素的 3×3 方形；(c)包含 21 个像素的 5×5 八边形区域；(d)包含 25 个像素的 5×5 方形；(e)包含 37 个像素的 7×7 八边形区域

中值滤波器也可用于降低前面图像平均操作中的随机或斑点噪声。图 4.18 显示了对图 4.9(a)所示原始图像应用具有 3 个和 7 个像素宽的邻域的中值滤波器的结果。与权重相乘比较，中值滤波器有两个主要的优点。首先，该方法不会降低或模糊各阶的亮度差异，因为可用值仅是那些出现在邻域中的

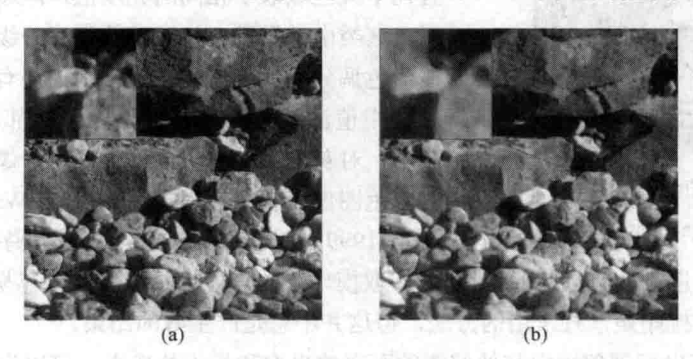


图 4.18 中值滤波器结果［与图 4.9(a)所示原始图像比较］：(a)半径为 1（3×3 像素的方形）邻域；(b)半径为 3（7 像素宽的八边形）邻域

值，而不是这些值的平均值；其次，与平均操作不同，中值滤波不会使得边界移位，它取决于邻域中出现的值的相对幅度和边界的局部曲率。克服这些问题后，中值滤波器无论是对于外观检查还是对于随后的图像测量都会是首选（Huang., 1979; Yang & Huang, 1981; Weiss, 2006）。

图 4.19 比较了对一幅带噪 SEM 图像应用高斯滤波和中值滤波后的结果。很明显，中值滤波器更好地保留了边缘锐度。即使对于非常大的邻域［见图 4.19(d)］，边缘也不会移位，但当邻域半径大于目前所呈现的任何特征的尺寸时，这些特征会被删除并被周围的背景值所取代。这是中值滤波器的一个基本特征，有时这可以更有效地从图像中去掉小的、不想要的特征。

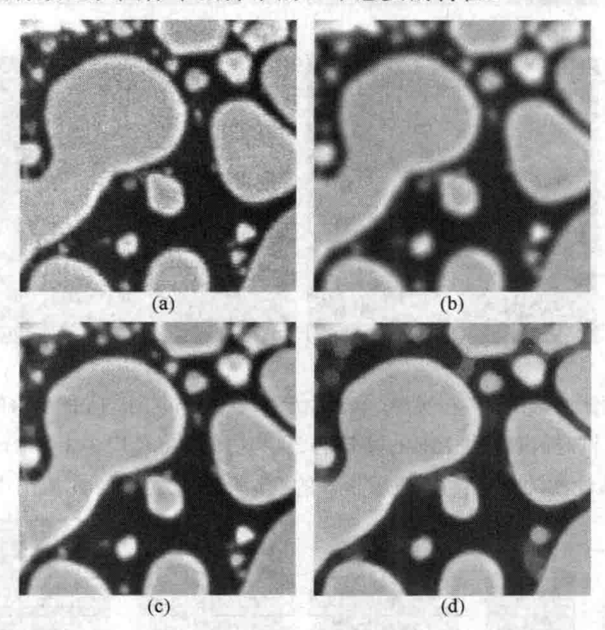


图 4.19　降噪技术比较（SEM 图像已放大到显示像素细节）：(a)原始图像；(b)标准差为 2 像素的高斯滤波后的结果；(c)半径为 2 像素（5×5 的八边形邻域）的中值滤波后的结果；(d)半径为 15 像素的中值滤波后的结果



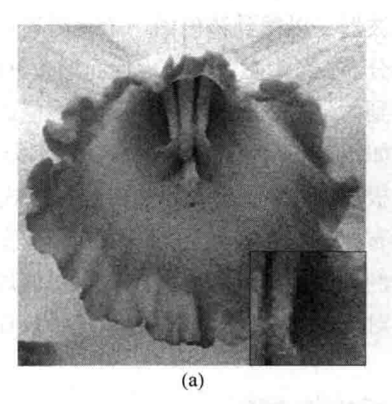
图 4.20　应用 3×3 中值滤波器 12 次后的结果［与图 4.9(a)所示原始图像比较］

因为中值滤波对边缘的退化最小，因此可以反复应用这一方法。图 4.20 显示了一个例子，它对一幅图像重复应用了一个 5×5 的八边形中值滤波器 12 次，直到像素不再出现变化（有时也称为“迭代中值滤波器”）。这一过程消除了细节，且几个大区域取了相同的亮度值，但边缘保持不变且轮廓分明。这类由于重复中值滤波的亮度分级，有时也称为画等高线或色调分离（但这些术语在别处还有其他含义）。

中值滤波器的概念需要一个对邻域内像素进行排列的顺序，对灰度图像而言该排列顺序由像素值提供。这一思路对彩色图像而言比较困难（Sartor & Weeks，2001；Comer & Delp 1999；Heijmans，1994）。将排序简单应用到红色、绿色和蓝色通道没什么用，而且因为色调以 360° 取模环绕的角，也不能将其很容易地推广到 HSI。仅按亮度而忽略颜色进行排序是一种常用的方法，但这并不总是产生好的结果。

将中值定义为其坐标到邻域中其他像素的距离的平方和最小的像素，可以设计一种彩色的中值滤波器（Astolo et al., 1990; Oistämö & Neuvo, 1990; Russ, 1995b; Celebi et al., 2007; Morillas et al., 2007;

Celebi & Aslondogan, 2008）。必须考虑测量这些坐标距离的彩色空间的选择。如前所述，对于处理而言，HSI 空间通常比 RGB 更受欢迎，但是对于诸如这些不同空间的相对缩放因子或如何处理的色调值的角度衡量等问题没有唯一的答案。图 4.21 显示了用于从图像中去除随机斑点噪声的彩色中值滤波器的一个例子。


(a)

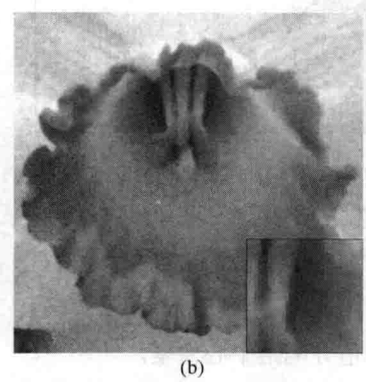

(b)

图 4.21　放大来显示噪声的数码相机图像的一部分：(a)原始图像；(b)应用彩色中值滤波器（半径为 2 像素）后的结果

将中值滤波（或排序滤波）从一个简单的标量数（灰度像素值）排序扩展到代表彩色空间值的矢量，可进行更为复杂的矢量排序。基于矢量角而非矢量长度来构建排序，可移除指向不正常方向的矢量（颜色），或构造结合方向和距离的混合方法（Smolka et al.，2001；Hanbury & Serra，2002；Jin et al.，2008）。矢量不必局限在三维，所以有可能将许多准则合并到像素相似性的一个全测度中。这些方法的性能以计算复杂度作为交换，但在某些情况下有可能以好的精度重构受到相当数量的斑点或散粒噪声污染的图像。

控制邻域的大小提供了定义要将被移除的缺陷大小的基本工具。在图 4.22 所示例子中，油画中已出现了细微的裂纹。对该图像应用平滑卷积，会混合裂缝的灰度值并模糊边缘。对该图像应用中值滤波则填充了这些裂纹，且不会模糊边缘，就像上面所示的从灰度图像中移除像素噪声一样。结果是恢复了原画的外观。

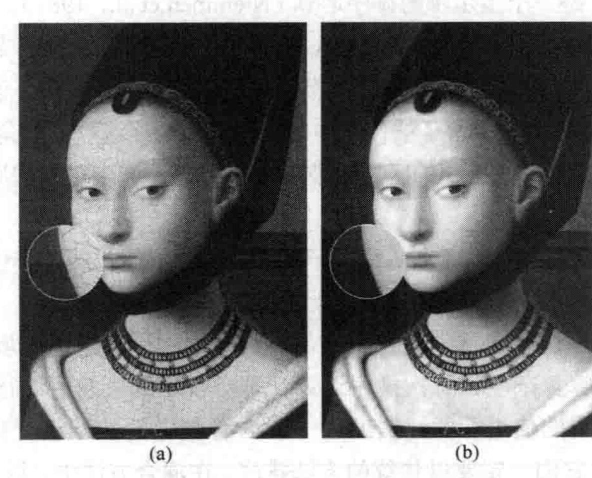

(a)　(b)

图 4.22　文艺复兴时期画家 Petrus Christus 于 1470 年绘制的“女士画像”：(a)呈现的外观，显示出裂纹和颜色的退化；(b)应用彩色中值滤波器填充暗色的裂纹

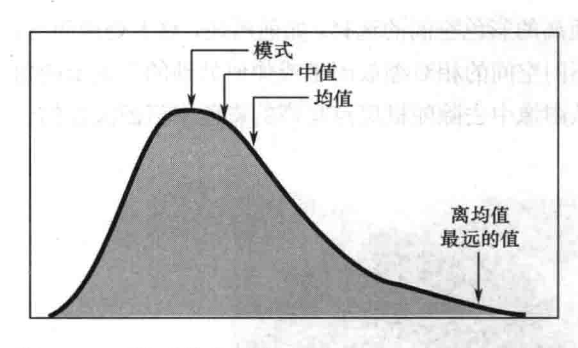


图 4.23 亮度值的非对称直方图分布，它显示了模式、中值和均值之间的关系。截尾中值滤波器通过丢弃分布中离平均值最远的值来操作，然后用剩余的中值作为对模式的一种估计

至少在原理上，使用一个滤波器可以更好地完成降噪，同时保留边缘的清晰度（Davies，1988）。根据定义，在每个邻域中亮度值的分布模式是最可能的值。但对于小邻域，模式很难定义。使用截尾中值滤波器可以得到该值的一个有用逼近。对于任何非对称分布，比如在大多数附近而非正好跨越边缘的位置得到的，模式是最高点，而且中值比均值更靠近模式，如图 4.23 所示。截尾中值技术会从邻域中丢弃一些极值，以使剩余像素的中值向模式移动。对于对称分布，值从两端丢弃，中值并不改变。在图 4.24 所示例子中，通过跳过亮度值与均值最不同的两个像素，对一个 3×3 的邻域进行了这种操作，对剩余的 7 个值进行排序，并将中值分配给中心像素。

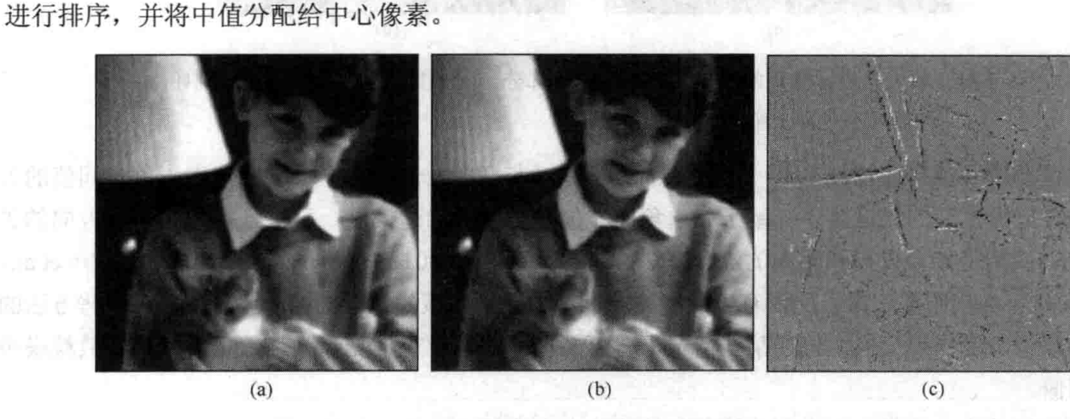


图 4.24 截尾中值滤波器的应用［原始图像在图 4.16(a)中显示］：(a)传统的 3×3 中值滤波器；(b)截尾中值；(c)图(a)与图(b)的差，对比度已扩展到显示沿边缘的值的差

中值滤波器的另一种改进用于克服其删除比邻域宽度的一半还要窄的线及将角变圆的趋势。所谓混合中值，或角保持中值，是一个多步骤的排序运算（Nieminen et al.，1987）。在一个 3×3 像素邻域中，像素可能在两个不同的组中排序，如图 4.25 所示。形成一个“X”形的 45° 邻域的中值以及形成一个“+”形的 90° 邻域的中值（两个组都包括中心像素）与中心像素进行比较，该组的中值然后保存为新像素值。如图 4.26 所示，这种方法保留了被传统中值滤滤删除的线或磨圆的角。在一个更大的邻域中，可以采用更多的方向；可以在一个 5×5 的区域中使用 4 个方向，产生可以同原中心像素值的三次重复一起排序的 4 个值。

混合中值也可以通过使用矢量距离来从每个子群中选择中值而扩展到彩色图像，如上所述。它可在保持清晰边缘和精细线条的同时，降低随机和散粒噪声，如图 4.27 所示。

混合中值涉及多次排序运算（首先是在每个像素组中，然后将这些中值与中心像素进行比较）的事实并不会明显增加计算复杂度。每次排序运算都是对较小数量的值进行的，而不是对同样大小的方形或八边形区域中所用的值进行的。例如，一个 5 像素宽的邻域包含有 25 个（方形邻域）或 21 个像素（八角形邻域），它们一定要以传统的方法排序。在混合方法中，每个子组只包含 9 个像素，而最终的比较只涉及 7 个值。即使对于值的其他逻辑操作，混合方法与传统的中值方法基本上也一样快。

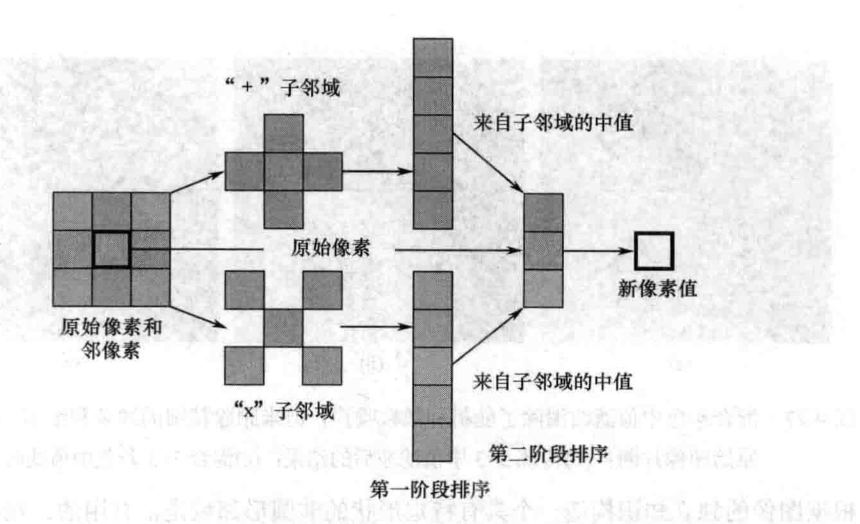


图 4.25　用于 3×3 混合中值滤波器的邻域像素图。两组都包含中心像素并分别进行排序。每组的中值及中心像素然后再次排序以选择最后的中值

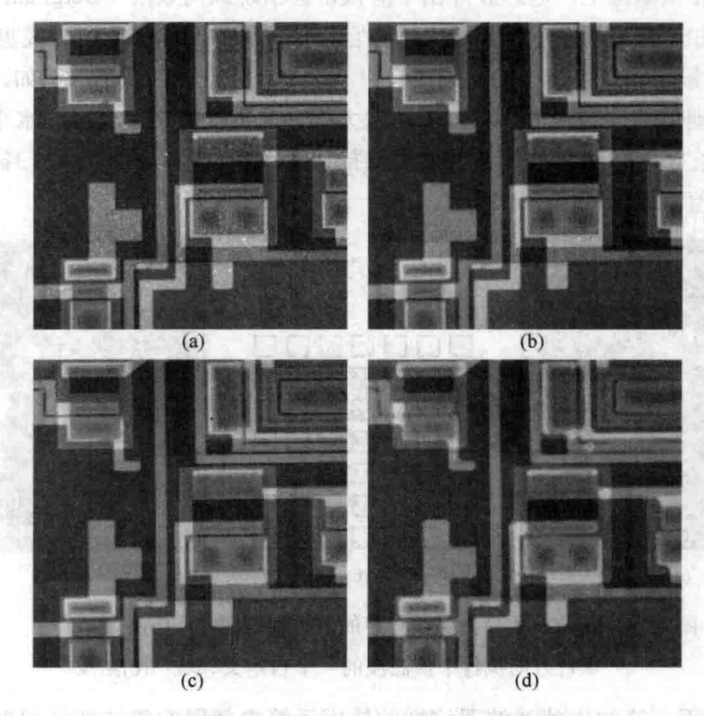


图 4.26　对集成电路应用混合中值滤波，增强了线条和角点：(a)电镀层中带有多余表面纹理的原始图像；(b)应用 5×5 混合中值滤波器后的结果；(c)应用 9×9 混合中值滤波器后的结果；(d)应用传统的 5×5 中值滤波器后的结果，它消除了多数纹理，但也使得角点变圆了，并填充了断线和裂隙

如果反复地应用混合中值滤波器，那么它也可以产生色调分离。因为线和角的细节由混合中值保持，与传统中值滤波相比，区域的形状没有那么平滑，尽管跨阶的亮度值仍然是锐利且色调分离的。对一幅图像色调分离，或减少灰度级的数量，以便区域变得在灰度值上均匀而且区域之间的边缘变得陡峭，更多地属于增强而不是纠正缺陷，但这里作为中值滤波的副作用提及。其他方法可以产生这种效果。例如，极值滤波器用邻域中的最小或最大值中最接近平均值的一个替换每个像素值。这个滤波器并不保持边缘，并且有可能使边界发生移位。当极值滤波器被迭代时，有时也称其为“切换”滤波器（Arce et al., 2000）。

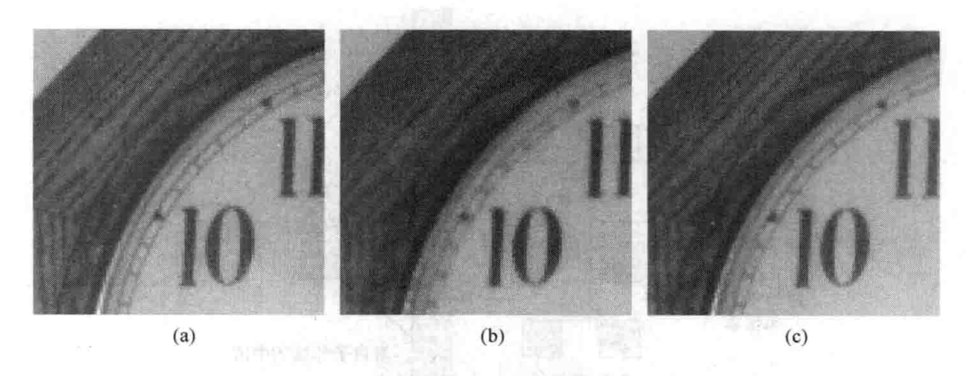


图 4.27 混合彩色中值滤波消除了随机和散粒噪声，但未删除精细的线条和细节：(a)放大的原始图像片断；(b)传统 3×3 中值滤波后的结果；(c)混合 3×3 彩色中值滤波后的结果

有时根据图像的独立知识构造一个具有特定形状的非圆形邻域是很有用的。在视频图像中的一种常见缺陷是出现水平线，这些水平线指出了两个隔行场之间的信号变化。由于直流信号偏移，类似的缺陷也出现在原子力显微图像上。电影胶片由于磨损常会出现纵向划痕（Bergman et al.，2008）。若水平线或划痕的方向已知，则可构造一个邻域来消除它。从电影中消除架空电力线也可以用这种方法。

图 4.28 显示了质量较差监控带中的视频线噪声。这是一种常见的缺陷。例如，它可能由脏录制磁头引起。垂向占支配地位的一个中值邻域可以消除大部分此类线噪声。类似的水平条纹经常出现在扫描显微图像中，原因在于对放大器时间常数而言扫描速率太高，因此模糊了沿扫描线方向的信息。对垂直条纹应用一个中值滤波器校正了该问题。

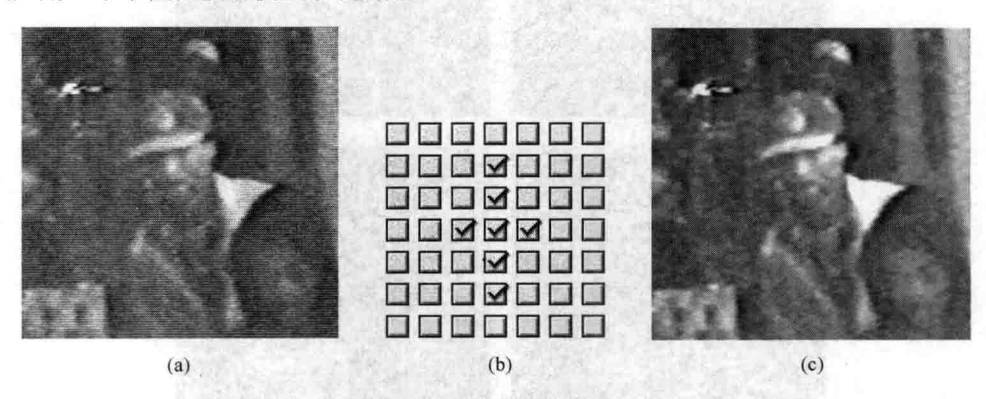


图 4.28 (a)带有水平扫描线噪声的监控视频图像；(b)主要在垂直方向执行中值滤波的一个自定义邻域；(c)结果

在多数简单情况下，诸如中值滤波器这样的排序运算中使用的邻域都是圆形，圆形的半径基于要保留的最小特征的尺寸来选择。为计算简单，有时会用一个方形邻域，但这可能会在结果中引入一些方向性的偏差。也可使用一个自适应的邻域，即它根据像素的值和/或位置来包含像素。条件或自适应选择始于一个更大的候选像素邻域，如果它们满足其他条件，就会包含在排序中，这可能会组合使用。若自适应邻域中的像素数是偶数而非奇数，则此时因没有中值存在，会面临几种选择。最可能的选择是对排序后列表中间的两个值求平均。在自适应邻域中选择像素的方法包括：

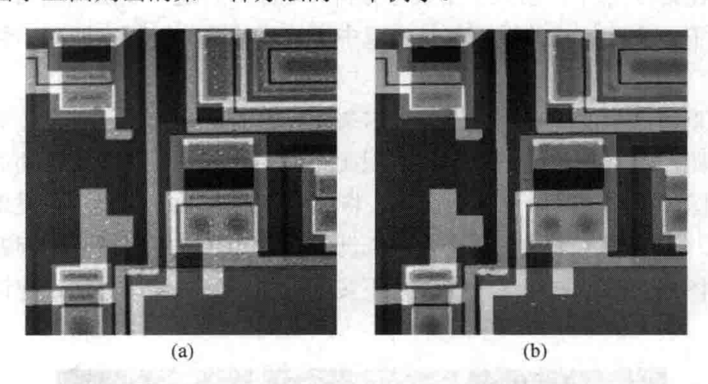
（1）与中心像素的差小于某个可调阈值的像素；

（2）在较大邻域内与中心像素值最接近的 N 个像素；

（3）在较大邻域内与中心像素值最接近的 N 个像素，这些像素与中心像素之间以及这些像素彼此之间是连续的（Kober et al., 2001）；

（4）根据它们到中心像素的距离对邻域中的像素进行加权，方法是多次将接近的像素值输入到将被排序的列表中。

以计算复杂度为代价，这些方法可良好地抑制冲激或散粒噪声，消除加性噪声，同时保持线和边缘不变。图 4.29 给出了上面列出的第一种方法的一个例子。



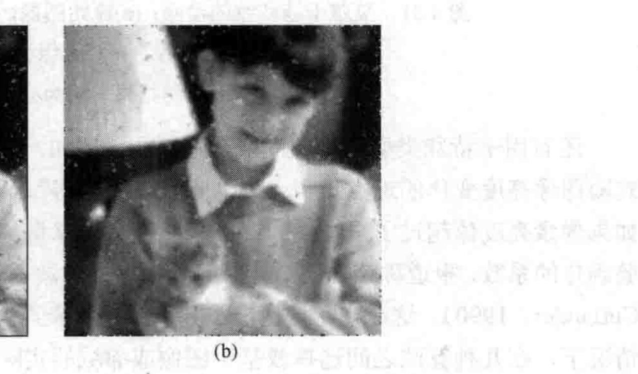
(a) (b)

图 4.29　使用自适应邻域的条件平滑：(a)原始图像［与图 4.26(a)同］；(b)9 像素宽的圆中的像素的中值，该圆在±40 个灰度级值范围内

4.5　其他邻域降噪方法

试图实现中值滤波器某些优点的对邻域值的简单平均的一个改进，是所谓的奥林匹克滤波器。这个名称源于奥林匹克运动会的一些项目的评分规则，即去掉最高分和最低分，然后对剩下的分数取平均。对邻域中的像素值也可进行同样的处理。去掉极端值，可以消除散粒噪声。剩下像素值的平均值然后用做新的亮度值。图 4.30 给出了对图 4.16(b)中的散粒噪声进行这种应用的结果。与中值滤波的结果相比，这种应用的结果要差一些：特征被模糊了，且噪声并未完全消除。





(a) (b)

图 4.30　对散粒噪声应用奥林匹克滤波器。原始图像如图 4.16(b)所示：(a)去掉了每个 3×3 邻域中的两个最亮和两个最暗的像素，剩下的 5 个像素值取平均；(b)去掉了每个 5×5 邻域中的 4 个最亮和 4 个最暗的像素，剩下的 17 个像素值取平均

除上述改进外，还存在其他形式的改进或条件平滑，这些改进会去掉来自平滑操作的移动邻域中的某些像素值。去掉某些像素值的原因总是基于这些像素不同于可用的像素，因此假定它们属区不同的区域。多数此类技术既会考虑区域中这种像素的位置，也会考虑其与原始中心像素的差。例如，对像素进行函数拟合，可检测斜率急剧变化并忽略位于边缘之外的像素（“克里金插值法”）。所有这些方法均根据图像和噪声的性质做了一些假设，假设越符合实际情况，结果就越好。

关于图像获取知识的一个特例，对于降低数码相机的噪声很有用。如第 1 章所示，典型的相机会记录单独的红、绿和蓝信号。硅检测器在频谱的蓝端最不敏感，而且典型的检测器设计会使用比绿色更少的蓝色过滤晶体管，导致蓝色通道比绿色通道出现更多的斑点噪声。因此，可以对不同的通道采用不同的降噪参数。在第 5 章中，引入了主成分分析，这是一种进一步从图像信息中分离出噪声的方法。在所有这些情形下，实际的滤波过程就是本章中所描述的那些过程，即基于邻域的加权平均法或基于排序的中值方法。

相机固有的抽样率所产生的另一个问题是莫尔条纹，如图 4.31 所示。外套上的图案间距接近于检测器中各传感器的间距，且有较小的角度差异，因此产生了随着人体的轻微移动而移位的干涉图案，这很让人分心。此类图像缺陷常见于电视图像中。将图像分离成 Lab 通道后，发现干涉图案几乎完全位于彩色通道而非 L（或亮度）通道。用一个高斯滤波器对图像进行平滑后，问题得到了解决，如图所示。当然，对电视图像实时地进行这一操作是不实际的（更好的解决办法是建议人们不要穿有花格或条纹的服装）。

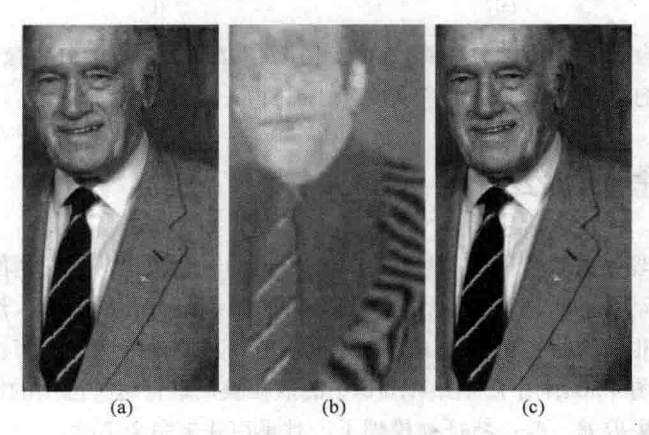
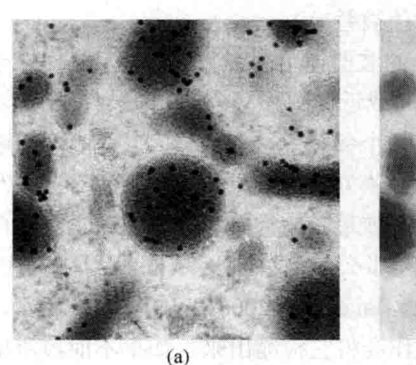


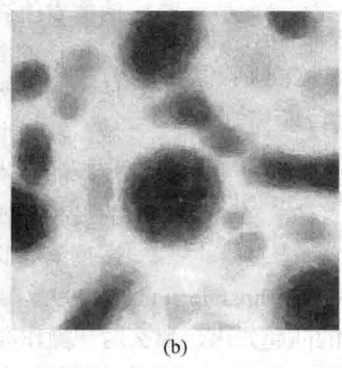
图 4.31　莫尔干涉条纹的消除：(a)原始图像，外套面料上显示有彩色的流苏花边；(b)转换到 Lab 空间后的 a（红-绿）通道；(c)在彩色通道中对流苏花边区域应用高斯平滑后的结果。因为亮度值未改变，图像的清晰度未受影响

还有用于特殊类型图像的复杂操作组合。例如，合成孔径雷达（SAR）图像包含有以一种已知方式随图像亮度变化的斑点噪声。为消除该噪声，将每个像素的亮度与一个局部邻域的平均值相比较。如果像素亮度值超过了平均值和标准差的某个计算量，它就会被一个加权平均值取代。使用一些由实验确定的系数，报道称该方法与一个简单的中值滤波器相比，在提高信噪比方面表现得更好（Nathan & Curlander，1990）。这是在特定情况下基于信号和噪声的特征知识的一个特定处理方法良好示例。一般情况下，在几种算法之间选择或基于图像或邻域的实际内容修改其算法的滤波方法，称为自适应或条件滤波（Mastin，1985）。

噪声通常建模为高斯加性函数，而降噪方法经常需要采用往图像中增加高斯噪声的方式来进行测试，但事实上各种噪声源具有非常不同的特点。光子或粒子计数产生的噪声一般呈泊松分布，当数字很大时会变成高斯分布。在胶片内曝光的颗粒的分布也近似为高斯分布。相机和放大器中的电子电路效应可以是加性的或乘性的，它对暗区与亮区的影响不同。大气或表面光散射造成的噪声通常是乘性而非加性的。雷达成像中遇到的斑点干涉更为复杂，因为它受相位移动和极化的相互作用。它通常也具有空间相关性，这就使得它更难建模。天文领域采用获取并组合持续时间很短的许多图像的方法来消除大气斑点，但这一内容超出了本书的范围。

靠排序进行滤波的另一种方法是，使用最大和最小亮度而非中值。这些操作通常称为灰度形态学运算，腐蚀和膨胀类似于在二值图像上进行的腐蚀和膨胀步骤，如第 8 章中所述（Heijmans，1991）。先腐蚀再膨胀的操作称为开运算，反之则称为闭运算。图 4.32 显示了开运算的结果，它在未引起结构形状失真的情况下，消除了图像中灰色细胞器和较亮细胞质上的暗色颗粒。这需要两次遍历图像，且在每次遍历期间，只使用先前的像素亮度值来生成新值。





(a) (b)

图 4.32 灰度开运算：(a)组织中金粒的原始 TEM 图像；(b)开运算后的结果。执行了两次排序运算：腐蚀用邻域（5×5 的八边形）中的最亮值来代替每个像素值；然后，通过使用该幅图像，膨胀用相同大小领域中的最暗值来代替像素值。去掉了暗色的小颗粒

作为一种消除噪声的方法，这一技术看起来完全不同于去掉极端值的中值滤波。不同在于，这种方法可选择性地消除亮（或暗）的人为缺陷或噪声，而中值滤波会将两者都消除。它可能有助于将图像可视化为一个表面，其中亮度表示高度，且这种方法对于诸如由扫描探针仪器得到的表面高程数据特别有用。图 4.33 给出了这种情形的一维表示。在第一次遍历时，每个区域中的最亮值被用于构造一条紧随“树顶”的新曲线。在第二次遍历时，每个区域中的最暗值被用来将曲线还原为那些大到足以在第一次遍历时幸存下来的点，生成一条忽略原图中暗噪声尖刺并同时保留明亮特征的新曲线。以其他顺序（先腐蚀再膨胀，称为闭运算）应用同样的运算，可移除明亮的噪声，同时保留暗特征。

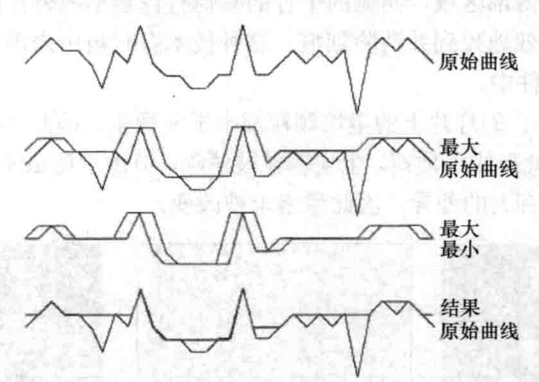


图 4.33 一维灰度腐蚀和膨胀运算示意图，它表明（从顶部开始）：原始曲线和第一次（最大）遍历的结果，产生一条通过最亮点的新曲线；步骤 2，生成一条通过步骤 1 结果中最暗（最小）点的新曲线；最终结果与原始曲线的比较，运算结果消除了噪声和暗色尖刺，但保留了明亮的特征

在两个大小不同的邻域使用排序的另一种方法，对于定位和消除噪声很有用。这种方法通常称为“顶帽”或“滚球”滤波器。假设像素的亮度值代表一个表面的高度。顶帽滤波器由放在该表面之上的

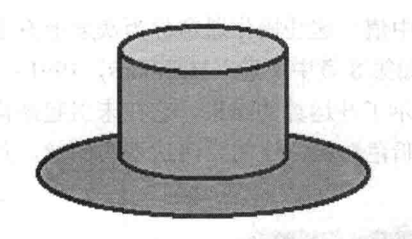


图 4.34　一个顶帽滤波器的图示。帽檐放在对应于图像中像素亮度的“表面”上。任何亮度值大于帽的凸起的像素将被检出，并被外部邻域的均值或中值代替。帽檐的内半径和外半径及冠顶的高都是可调参数

一个平圆盘和一个直径较小的中间凸起组成，如图 4.34 所示。该滤波器以图像中的每个像素为中心，将帽檐“靠在”表面上。通过帽的凸起的任何“突出”像素均被认为是噪声，并被替换。替换值可能是帽檐所覆盖的像素的平均值或中值（“滚球”一词源于假设斑点表示凹陷，而表面之上滚动的球所接触到的孔底的像素值被替换）。

该滤波器使用两个邻域，一个是对应于帽子内部或凸起区域的圆（或近似圆形），第二个是对应于边沿并围绕它的环形邻域。在每个邻域中，找到最大（最亮或最暗，具体取决于将被消除的噪声）值。如果内部“凸起”区域和外部“边沿”区域的最亮（或最暗）像素之间的差超过一个门限值（帽子凸起的高度），那么就用外部区域的平均值或中值替换这个最大值（Bright & Steel，1987；Deshpande et al.，1999；Bai et al.，2008）。如图 4.35 所示，如果一个特征完全位于较小的邻域（帽的凸起）内，那么这一操作会将它检测出来。如果不能检测出，或者各特征间的间距小于帽边沿的宽度，则内部和外部间没有大的差异，就什么也检测不到。

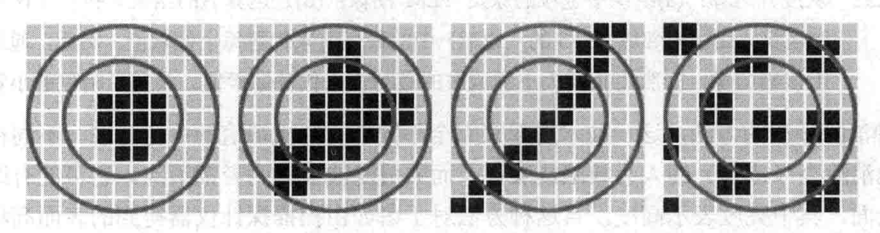
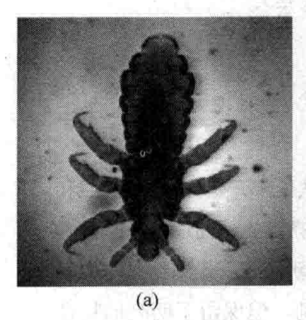


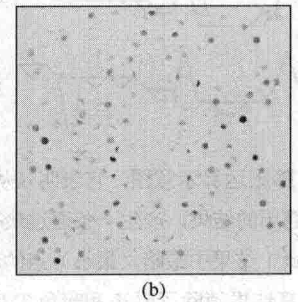
图 4.35　顶帽滤波的运算：如第一个例子所示，检测出了完全包含在内部区域（“冠顶”）中的所有特征。大过冠顶或相隔不超过帽檐宽度的特征未被检测出

用顶帽或滚球找到特征后，可将它们删除，方法通常是使用由边沿内插得到的值来替换它们。对沟槽形邻域（一行像素形成内部区域，两侧的平行的其他行像素形成外部区域）应用这种方法，在已知划痕方向的前提下，可有效地找到并消除划痕。这种技术有时被用来消除数字电影中的划痕，且被内置于一些胶片扫描仪的固件中。

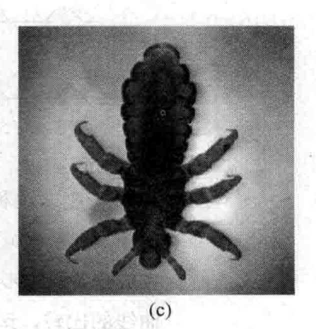
图 4.36 显示了一个例子。幻灯片上的尘埃颗粒都小于 9 像素，因此半径为 4 像素的内环和半径为 6 像素的外环组成的滤波器就形成了顶帽。尘埃颗粒被消除，但由于昆虫本身大到放不进内环中，导致两个邻域中最暗的值之间没有大的差异，因此像素未被改变。



(a)



(b)



(c)

图 4.36　顶帽滤波应用：(a)昆虫幻灯片上含有尘埃颗粒的原始图像；(b)通过顶帽滤波隔离尘埃颗粒；(c)使用帽檐内插的值替换冠顶中较深的值，消除了尘埃颗粒

接下来的章节介绍该滤波器对于不同目的的用法。替代将极端点作为噪声或灰尘消除，相同的运算可用来找到并保留比帽冠小且比局部邻域更亮或更暗的点。图 4.36(b)说明了顶帽的这一功能。通过对邻域大小使用同样的设置，但保留极端值并抑制那些差不大于顶帽门限值的值，恰好显示出了幻灯片上的尘埃。

4.6 缺陷消除、最大熵和最大似然

如上所示，滚球/顶帽滤波器可用于消除小缺陷，方法是使用取值周围邻域的值来替换像素。内插值是一种通用工具，它有时被用来替换局部缺陷，这些缺陷要么来自标本本身，要么来自相机（镜头上的灰尘）。若缺陷较大和/或形状不规则，则平滑后的图像中可能会留下一些来自缺陷的像素值，而对基于排序的滤波使用大到足以包围缺陷的邻域并不实际。在许多情形下，使用邻近的值来填充缺陷区域会得到更好的结果。在图 4.37 所示例子中，人为选择了一片树叶上的一个洞，且像素被轮廓边缘像素的线性插值取代。该算法效果上近似于拉伸一个膜，使其横跨一个不规则且边缘具有可变高度的开口。弹性膜形成一个能量最小的表面。以同样的方式，完成了区域内的像素值的计算。

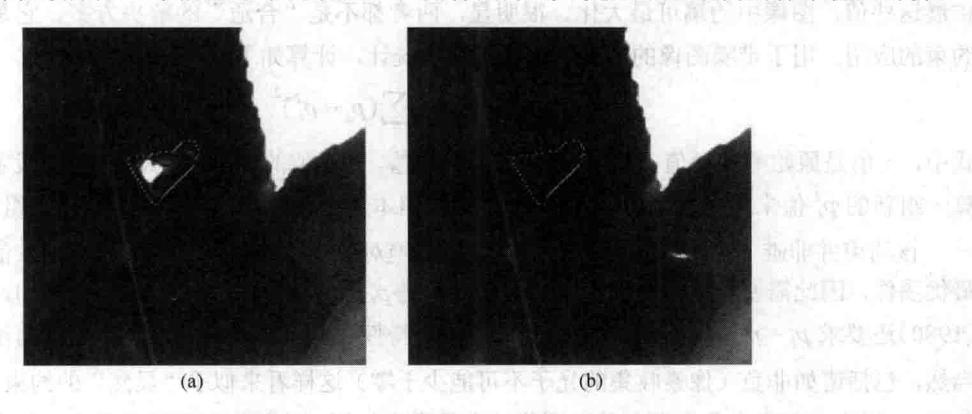


图 4.37　由区域的边缘进行内插来消除：(a)枫叶的放大图像，洞周围显示了人为绘制的区域；(b)由边缘内插的值

不采用人工画出轮廓，使用第 7 章所示的任何阈值处理技术也可选择区域。例如，指定可以从周围环境中辨别出缺陷的亮度范围或颜色值，或使用一种区域生长法，即在缺陷中选取一个点，进而允许软件包含所有相连的类似像素。当环境验证合理时，这种方法可用于从图像中消除人为选择的特征。但要注意，在科学成像情形下，向一幅图像中添加任何内容（如替换或覆盖已被删除的特征）是不恰当的。

除了噪声之外，图像还可能包含其他缺陷，如由运动或光圈散焦导致的模糊。第 6 章在解卷积主题下描述的“逆滤波”方法易受噪声影响。当出现的噪声本身可被表征时，就有可能使用最大熵方法消除人为缺陷。

根据图像内容的一些先验知识，以及关于将被清除退化的性质与复原图像的一些假设（Skilling，1986；Frieden，1988），基于熵的计算机密集型方法，可消除如噪声或图像模糊之类的人为缺陷。该方法的常规描述是，假设有一幅含有 N 个像素的图像，它共由 M 个光子形成（M 通常远大于 N）。任何单个像素（即像素的亮度值）中的光子数记为 P_i，其中 i 是像素序号。

观测图像已归一化到像素亮度值

$$p_i = \frac{P_i}{\sum P_i} \tag{4.4}$$

在没有人为缺陷时，归一化值只逼近可被收集的“真实”图像。p_i 值恰好是图像直方图。在约束条件下，用于逼近这种理想图像的方法，会更改亮像素亮度值，以便最大化图像中的熵。这种方法的证明由统计概率和贝叶斯定理给出，此处不再推导。在有些情形下，这种方法可显著增强图像的质量。

通过计算由排列组合所形成模式的方式的数量，即

$$S = \frac{M!}{P_1!/P_2!,\cdots,P_N!} \tag{4.5}$$

可正式给出亮度模式的“熵”，其中!表示阶乘。对于较大的 M 值，由斯特林（Stirling）近似公式可将上式简化为更为熟悉的形式：

$$S = -\sum p_i \log p_i \tag{4.6}$$

这与统计力学中熵的计算相同。根据信息论，在以 2 为底的对数情形下，图像的熵是用于表示图像所需的最小数量的位数/像素。

令每个像素的亮度为平均亮度，图像中的熵在绝对意义上可最小化，而沿着整个亮度范围均匀地扩展这些值，图像中的熵可最大化。很明显，两者都不是“合适”的解决方案。它是产生可用结果的约束的应用。用于带噪图像的常见约束基于卡方统计，计算如下：

$$\chi^2 = \frac{1}{\sigma^2}\sum (p_i - p_i^2)^2 \tag{4.7}$$

式中，p_i 值是原始的像素值，p_i^2 是改变后的亮度值，σ 是值的标准差。对 χ^2 的值设置一个上限，可计算一组新的 p_i^2 值来最大化熵。χ^2 的一个典型（但本质上任意的）极限是 N，即数组中的像素数量。

该约束并非唯一可能的选择。也可以选择差绝对值之总或其他一些加权规则。信息并不足以产生最优图像，因此需要添加其他的约束。约束之一是 p_i 与 p_i^2 值的总数必须相等。例如，Bryan 和 Skilling（1980）还要求 $p_i - p_i^2$ 值对应于成像源的期望噪声特性(例如，对于简单的计数统计为泊松或高斯分布)。当然，包括诸如非负（像素收集的光子不可能少于零）这样看来似乎“显然”的约束也很重要。Jaynes（1985）指出，实际上总会存在关于图像的大量真实知识，这些知识可用做约束，但它们因过于明显而被忽略。

对 p_i^2 值迭代求解，可得具有期望平滑度和噪声特性的新图像，新图像与原始图像相比，噪声特性会得到增强。最大熵方法的其他公式通过计算互熵

$$S = -\Sigma p_i \log\left(\frac{p_i}{q_i}\right) \tag{4.8}$$

而非总熵，可将一次迭代的图像与下一次迭代的图像进行比较。上式中，q_i 对于相同像素是先前的图像亮度值，对于理论图像是建模后的亮度。公式中，互熵将最小化。基本原理保持不变。

一幅图像的最大似然重建是另一种相关的方法，该方法假定区域间的边界是清晰的。在多数真实图像中，跨越任意边界的像素，会按照边缘相对于有限大小像素的精确位置，对边缘两侧的值取平均。这将导致看起来模糊边界的中间亮度或颜色值。将像素重新分类到一个或另一个区域的亮度或颜色，会锐化边界。

对于分类而言，最常见的统计方法之一是测量每个像素周围邻域内的许多子区中的像素的方差，如图 4.38 所示。具有最低方差的区域被用来表示中心像素应属于的区域，并被赋予该子区域的均值（Kuwahara et al.，1976）。图 4.39 显示了该方法的一个应用。邻域和子区域的大小对于更宽的边界可以更大。

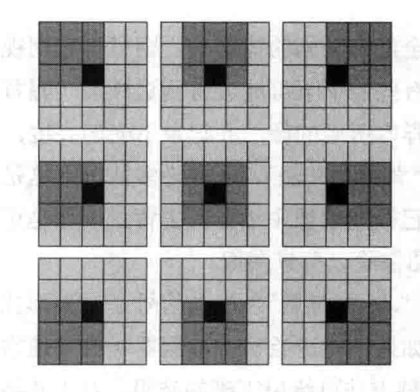


图 4.38 一个 5×5 邻域内的 9 个 3×3 子区域。Kuwahara 滤波器将具有最小方差区域的均值分配给中心像素

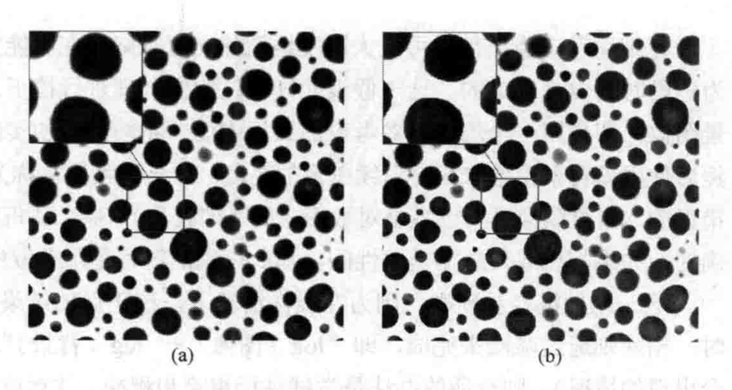


图 4.39 最大似然滤波应用：(a)泡沫中气泡的原始图像，边缘因为半透明而不清晰；(b)滤波后的结果，区域边界很清晰

4.7 非均匀光照

图像分析最常见的策略使用图像中各区域的像素值作为一种识别手段：无论在视野中的何处出现，假定同类特征都有相同的亮度或颜色。如果该亮度与图像中其他区域的不同，或通过适当的图像处理变得如此，如第 5 章所示，那么它可以用来为计数、测量或识别目的而区分物体或结构。即使存在基于亮度或颜色无法识别的其他类型的物体，后续的测量也能找到感兴趣的物体。

这种方法并非没有缺陷，这将在第 7 章介绍将连续色调灰度或彩色图像转换为二值（黑白）图像时，一并深入讨论。还存在其他方法，如区域生长或分裂与合并法，这些方法对均匀光照没有严格要求。但在使用这些方法前，记住简单的亮度阈值处理仍是迄今为止最易和最快隔离图像中各个特征的方法，因此考虑图像的明暗很重要。

在观看不规则表面时，即使表面材料和光度是均匀的，但每个区域散射到观看者或相机的光量，也和光源与观察者间的距离及方向有关。事实上，该原理可用于估计表面方向，所用技术称为从明暗恢复形状。人类视觉似乎很容易和迅速地运用这些规则，因为人们通常不会被真实物体的图像混淆。

本书涉及的大部分技术本质上是二维图像处理。无论图像是来自光/电子显微镜、宏观成像，还是来自卫星图像，表面高度的变化通常都会比横向尺寸小，这就是人们常说的“2.5 维”。当然，情况并不总是这样（考虑在 SEM 中仔细观看织物或在森林中树木的一幅照片），但这些问题可当做普通规则的例外情况，并意识到可能需要更为精心的处理。

浅起伏表面甚至不必是平的，一个例子是从卫星观看的地球曲率。这会在整个视野产生明暗变化。来自一侧的宏观或微观表面的光照也是如此。即使是精心收集的光或环形灯，也只能逼近场景的均匀光照。摄影棚中所用的柔光镜和遮阳伞是有用的，但这种情形下配置它们的目的，通常是作为备用，而不是作为均匀的光照。

对于透射显微成像，具有聚光透镜系统的均匀光源可以实现得很好，但这些系统也很容易失准而产生明暗。最后，第 1 章提到了镜头或相机可能会引起虚光，即图像的四角要暗于中心，原因在于部分光被光学器件中更大的通常长度所吸收，而改变镜头的有效孔径会减轻这种现象。也存在相机传感器本身的影响，如在使用广角镜头时非垂直入射的影响。通过记录一幅“背景”图像，可以校正所有这些影响（这种方法也可用于消除或降低第 1 章中提及的固定模式噪声）。

仔细调节成像条件，可最大限度地消除光照缺陷，在不能完全消除缺陷的情形下，则可将它们视为一段时间内是恒定的。这一假设允许通过图像处理进行校正。有些情形下可能会得到这样的一幅背景图像，即使用一个均匀的参考表面或标本而非将被成像的实际样本所成的像，并记录下光的亮度。然后使用该背景图像来衡量后续图像的亮度。这一过程通常称为“背景相减法”，但在许多情况下这是错误的。如果图像采集设备是对数形式的（如胶片），那么从每幅已获取图像中逐点减去背景图像是正确的；如果相机或传感器是线性的，那么正确的过程是用获取的图像除以背景图像。

这一差异很容易理解，因为在线性情况下，校正后的结果是“图像/背景”，而当信号是对数形式时，相除则通过减法来完成，即“log（图像）－ log（背景）”。如果不知道相机的检测器响应（通常会出现的情况），则合适的办法是尝试进行相除和相减，并观察哪种方法可给出正确的结果。对于其他传感器响应函数，没有完全正确的算术方法，必须首先校准响应，并用它来将测量的信号转换到线性或对数空间。

在两幅图像相减或相除的处理中，会丢失原始数据的一些动态范围。背景亮度的变化越大，结果图像中所保留的这种亮度级的变化就越小。在采取处理方法前，这一损失和两信号相减所致的不可避免的统计噪声的增加，要求首先采取所有实际的步骤来使得光照均匀并降低获取图像时的噪声，

第 1 章中的图 1.76 显示了这种调整的一个例子，其中可以分别获取背景照明函数。这一获取通常是通过从光路中拿开标本来完成的，例如用空白玻片取代显微镜载玻片，并存储一幅表示变化的图像。所存储的图像然后可用于进行调整。图 4.40 显示了一个使用翻拍台的例子。台基涂上了一种单调且均匀的灰色，因此移走物体就为从一边记录下光照图像提供了机会，而且移走物体可有效地调整原始图像。在该例中，相机按对数记录图像，因此从原始图像中逐点减去了背景图像。

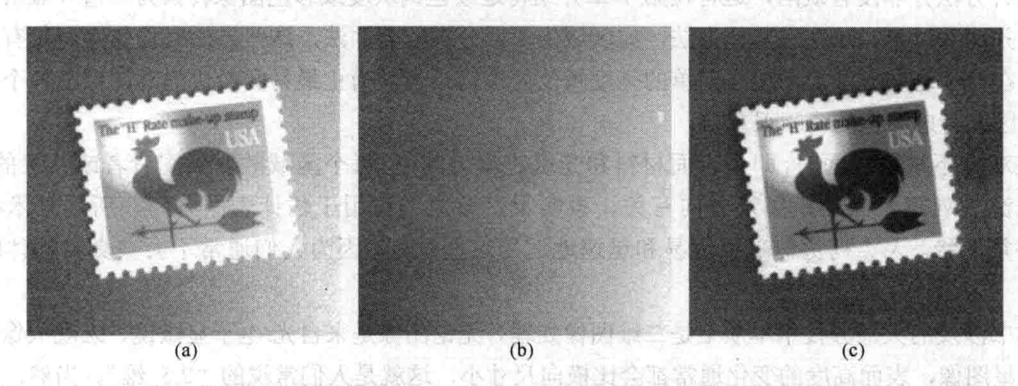


图 4.40 消除测量背景来平衡图像亮度：(a)翻拍台上物体不均匀光照的原始图像；(b)移走物体并记录背景；(c)逐点减去背景生成均匀的结果

图 4.41 显示了亮度变化可直接从原始图像减去的一种情形。通过分离图像中的红色、绿色和蓝色通道，可以看到绿色图像中包含了重要的细节：眼睛的视网膜血管，由于亮度变化它之前难以看到。红色图像几乎没有细节，但记录了一般的变化，因此绿色与红色之比（即除以背景）调匀了变化的背景，并使得血管在对比度上更为均衡。

遥感图像通常使用不同颜色通道的亮度之比来降低表面形状效应（第 1 章中的图 1.60 给出了一个例子）。直接测量两幅图像的另一种情形，允许使用相除来校正出现在透射电子显微镜中的不均匀。这常用于厚度出现变化的标本中，尤其是材料科学中的无机样本。计算传统明亮视野图像与连到显微镜上的分光仪所分离的零损失电子之比，可直接没量标本的厚度变化，进而使用厚度变化来进行浓度和结构测量。

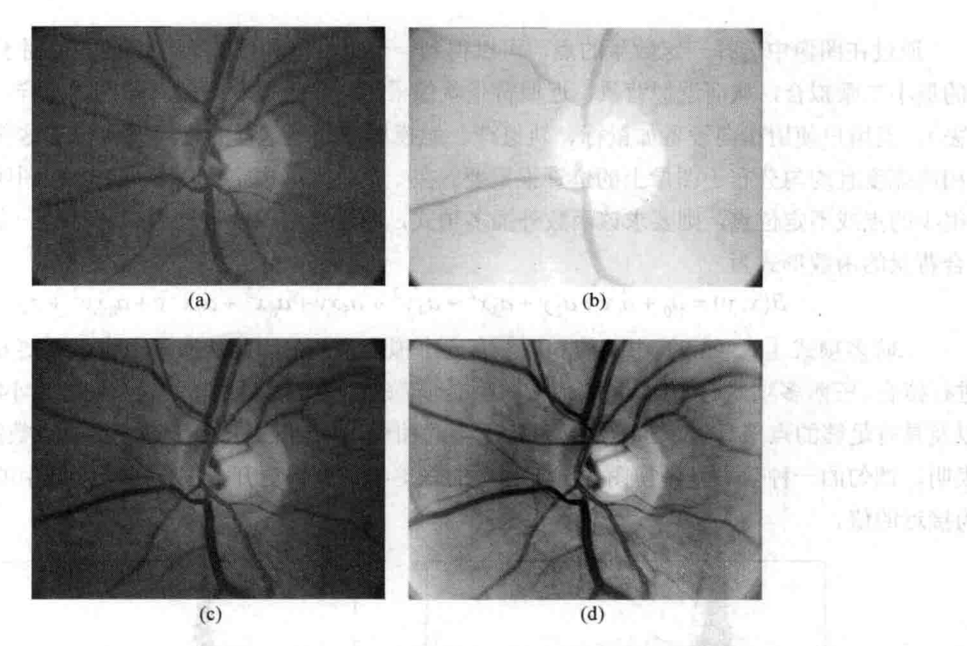


图 4.41　用观测的背景调匀图像：(a)眼睛的视网膜图像，其中毛细血管难以分辨；(b)红色通道细节很少明暗对比明显；(c)绿色通道对比最为强烈；(d)绿色通道与红色通道之比结果的对比度更为均匀

4.8　拟合背景函数

许多情况下并不记录背景图像。移除标本可能并不实际，其存在可能是亮度变化一个因素。这包括透射光、X 射线或电子图像，其中的标本厚度会变化，因此整个吸收会受到影响。另一种情况是，被检查的表面是倾斜的、弯曲的或不规则的，这会使得图像显示变化的背景。当然，有时这一步会在图像获取时忽略。使用图 4.40 的例子，如果背景图像未被记录，那么由内插来生成背景也是可能的。

图 4.37 所示的边缘内插方法，只使用了边界附近的像素值。这里，我们采用一个使用所有可用背景位置的多项式函数来定义背景亮度。对于与偏离中心的光学或照明相关的亮度渐变，多项式通常是较好的模型。从原始图像中的每个点删除计算出的亮度值，调匀了亮度，如图 4.42 所示。只要可以定位有代表性的背景（或任何具有一致亮度的结构）块，这种方法就可提供对图像区域的充分覆盖。

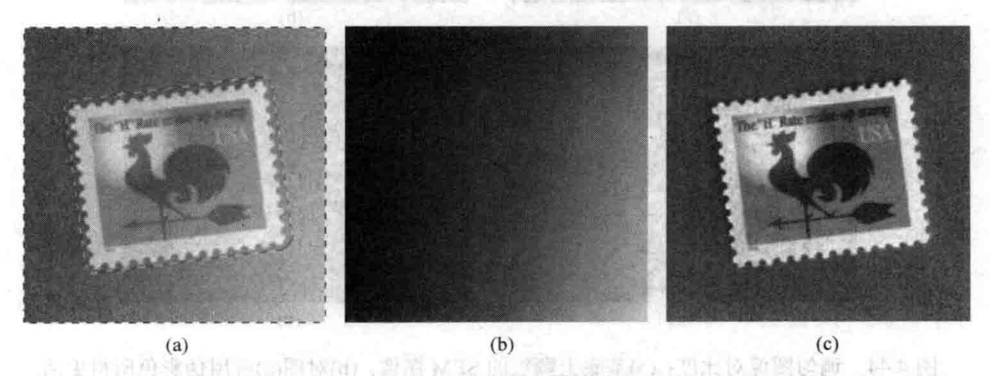
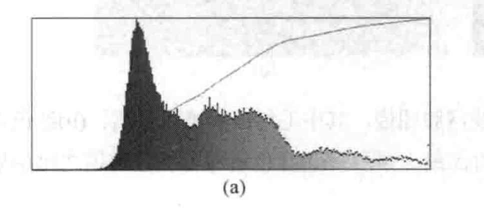


图 4.42　通过拟合调匀图像亮度：(a)在图 4.40(a)中选择邮票周围的一个背景区域（红色叠加显示屏蔽掉的区域）；(b)平滑多项式亮度函数拟合到背景点；(c)从整个图像中减去背景函数后的均匀结果

通过在图像中选择一定数量的点，可以得到一个亮度值和位置列表。这些值用于执行函数 $B(x, y)$ 的最小二乘拟合，从而近似背景，近似背景就像采用物理手段获得的背景图像那样，可被减去（或除去）。当用户使用指向设备如鼠标、轨迹球、触摸屏或光笔这样的指向设备标记这些点时，选取具有相同亮度且均匀分布于图像上的位置很重要。在一个小区域内定位许多点，而在图像的其他部分定位很少的点或不定位点，则要求该函数外插多项式，而这会引入明显的错误。对于一个三阶多项式，拟合背景的函数形式为

$$B(x,y)=a_0+a_1x+a_2y+a_3x^2+a_4y^2+a_5xy+a_6x^3+a_7x^2y+a_8xy^2+a_9y^3 \tag{4.9}$$

二阶多项式［式（4.9）中的前 6 项］有 6 个拟合常数，因此从原理上讲只需要 6 个标记点就可以进行拟合。三阶多项式有 10 个系数。但为了得到较好的拟合效果并降低对个别像素小幅波动的敏感性，以及具有足够的点来对整个图像区域进行适当的抽样，通常需要几倍于这一最小点数的标记点。图 4.43 表明，调匀的一种效果是使得图像直方图中的峰更窄、分得更开，因此每个区域中的所有像素都有更为接近的值。


(a)

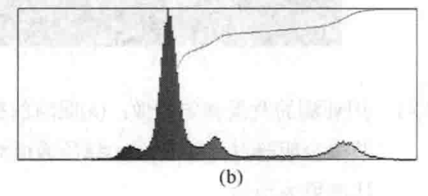
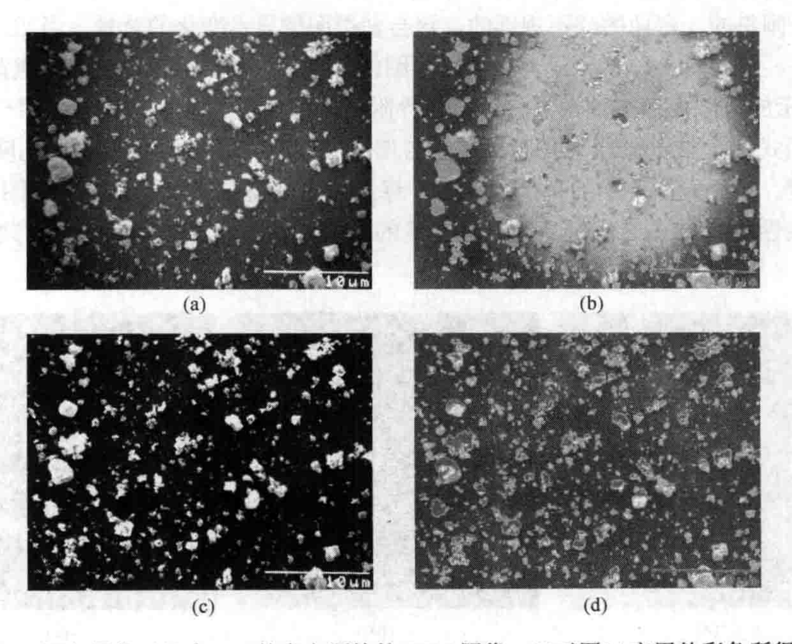

(b)

图 4.43　图像直方图调匀的效果：(a)图 4.40(a)的直方图；(b)图 4.40(c)的直方图

图 4.44 显示了一个例子，其中背景点是自动而非人为选取的。这幅 SEM 图像显示出了基底上的颗粒，原因在于表面和 SEM 电镜室的几何形状使得图像的边缘和角落看上去要比其他部分暗。如第 2 章所述，人眼适应亮度的渐变，因此使用伪彩色查找表来揭示图像中的阴影有时是有帮助的，如例中所示。



(a) (b) (c) (d)

图 4.44　调匀图像对比度：(a)基底上颗粒的 SEM 图像；(b)对图(a)应用伪彩色所得更清晰的阴影效果；(c)多项式拟后调匀后的图像；(d)对图(c)应用伪彩色后的效果

背景亮度的变化是位置的平滑函数，它可用一个多项式来逼近。这个例子中，颗粒要亮于局部背景，而靠近图像边缘的颗粒则要暗于图像中心的颗粒。利用背景由局部暗像素组成及其应均匀这一知识，函数的算法自动将图像划分为多个区域（例子中显示的9×9 网格），并找到了每个区域中的最暗点。然后使用这 81 个值及它们的位置，来生成式（4.9）所示多项式，再后将它从原始图像中移除。这就完成了图像亮度的调匀。

若有一个分布于整个图像区域的明显结构或相位且它包含最暗（或最亮）的像素，则自动调匀最简单。图 4.45 所示例子中的特征（毛穴）在明亮背景上呈暗色。由于表面曲率，标本（抛光的陶瓷制品）具有亮度的全部变化。每个标本区域中的最亮像素表示矩阵，因此都应相同。该例中，使用了相同的图像子网来选取最亮的局部值，然后采用最小二乘法，由这些局部值计算二阶多项式（6 个系数）。此时，拟合程序指出，在图像像素的整个 0～255 范围内，有小于 2 个亮度值的拟合误差（均方根值）。图 4.45(b)显示了用 $B(x, y)$函数计算出的亮度，图 4.45(c)显示了从原始图像中逐像素移除背景后的图像调匀结果。这一调匀操作移除了背景亮度的变化，并允许设置亮度阈值来勾画出毛穴用于测量，如第 7 章所述。

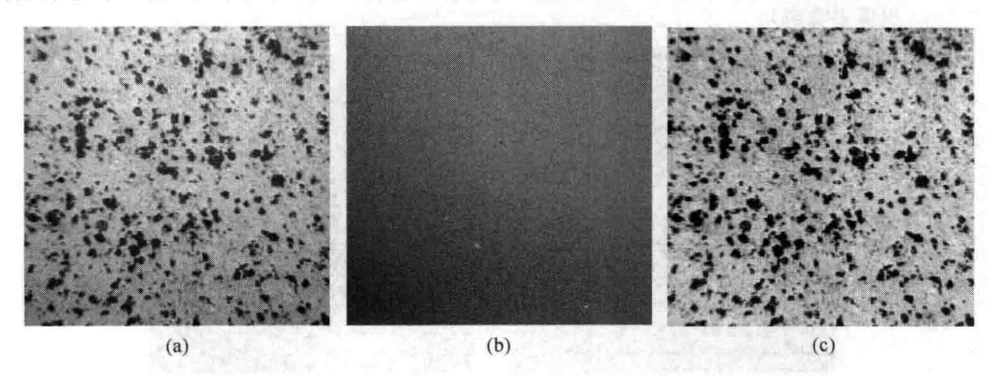


图 4.45 非均匀光照的自动调匀：(a)陶瓷标本的反射光显微图像，不均匀背景亮度由不平的表面引起；(b)拟合 81 个方形（9×9 风格）中最亮点的多项式所计算的背景函数；(c)从图(a)中减去图(b)后的调匀图像

如上面的例子所示，这种自动调匀的方法既适用于亮背景，也适用于暗背景。将这种方法同时应用于图像内每个区域中的最亮和最暗像素，就可以画出图像对比度与位置的关系函数，图 4.46 给出了亮度剖面线的一个例子。当图像由于不均匀光照或标本厚度变化而丧失对比度时，这种自动对比度调整方法非常有效。图 4.47 和图 4.48 给出了一些例子。在图 4.47 中，对比度变化是由标本厚度变化产生的，而在图 4.48 中，对比度变化则是由表面曲率产生的。在这两个例子中，去除阴影可使得整个对比度增大到显示更多的图像细节。第 5 章将介绍实现自适应对比增强的其他方法，这些方法会因图像中不同位置的亮度变化而进行不同的调整。

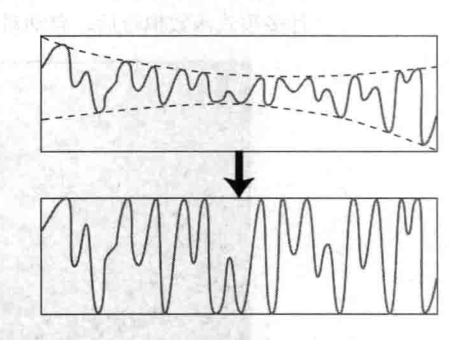


图 4.46 对整个图像中的最亮值和最暗值进行多项式拟合并在两个极值间进行亮度值拉伸，所得的对比度自动调整草图

有时用于消除整个亮度平缓变化的另一种方法，会采用第 6 章中讨论的频率变换。它假定图像中的背景变化是低频信号，且可在频率空间中从定义所示特征的高频区分出来。当这一假设合理且这些频率对应于可被识别的背景时，在频率空间表示中使用一个简单的滤波器就可将它们消除。

图 4.49 给出了这种方法的一个例子。原始图像中的亮度变化由显微镜中偏离中心的光照引起。使用二维傅里叶变换将图像变换到频率空间（详见第 6 章），并对频率空间图像滤波后，降低了前四个频率分量的幅度，尔后的反变换就生成了所示的结果。

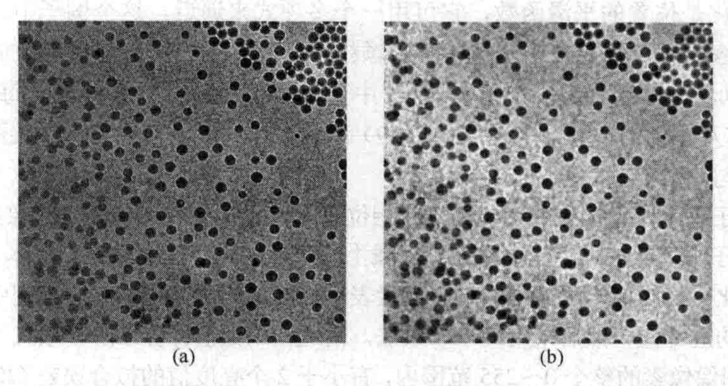


图 4.47 乳胶颗粒的 TEM 图像：(a)标本厚度变化所至对比度变化的原始图像；(b)对亮和暗像素值进行多项式函数拟合后，自动对比度调整后的效果（已应用一个中值滤波器来降低随机像素噪声）

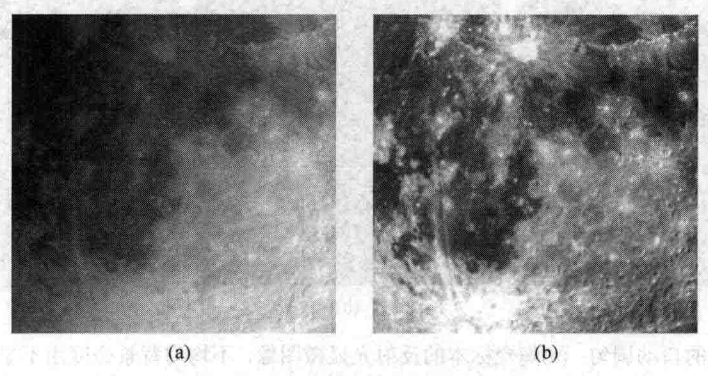


图 4.48 月球表面的望远镜图像：(a)表面曲率所致对比度变化的原始图像；(b)对亮和暗像素值进行多项式函数拟合后，自动对比度调整后的效果

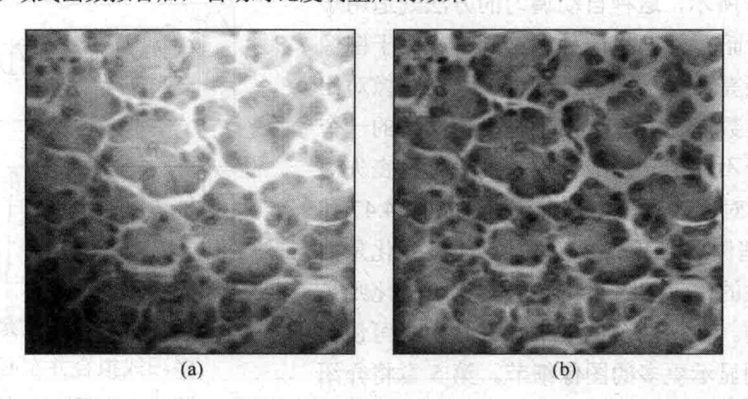


图 4.49 删除二维傅里叶变换中的低频项后，图像亮度和对比度的调匀结果：(a)非均匀光照原始图像；(b)将最低频率分量的幅度降为零后的亮度调匀效果。注意，除边框附近的变化外，类似结构的亮度在整个图像中并不恒定

对于这幅图像，这一方法并不完全成功。由于频率变换尝试匹配的左边缘和右边缘以及顶部边缘和底部边缘，图像边缘存在明显的变化。此外，整个图像中有着相同外观且期望恰好具有相同亮度的暗和亮区域的亮度存在局部变化，因为亮度是局部细节（包括实际亮度、间距和特征形状）的函数。实际中，这种调匀方法通常不能得到令人满意的结果。

4.9 排序调匀

当背景变化剧烈而不能用简单的函数拟合时，可用一种不同的方法。当表面不规则（如 SEM 图像中裂面的细节、日光照射的山脊）时，这种方法尤其有用。这种方法的基本假设是，感兴趣特征尺寸有限且至少在某个维度上小于背景变化尺寸，而且背景在任何位置要么亮于局部特征，要么暗于局部特征。实际中这两种要求通常会得到满足。

本章前面介绍了邻域排序运算（如中值滤波器），第 5 章中还将会用到这一运算。邻域排序运算的基本思想是，在一些小区域（通常为尺寸可调的近似圆形）中，将每个像素与其邻域进行比较。对图像中的每个像素均执行该运算，所得结果是其像素被邻域值替换的一幅新图像。

对于背景调匀，如果已知背景局部暗于特征，则用邻域中最暗的像素值代替原像素值。当背景亮于特征时，则用邻域中的最亮像素值代替原像素值。如上所述，这些运算有时称为灰度腐蚀和膨胀，类似于第 8 章中讨论的适用于二值（黑白）图像的形态学运算。对整幅图像应用这一操作的结果是，按邻域区域的半径收缩特征，并将局部背景亮度值扩展到之前被特征覆盖的区域。

图 4.50 对一幅暗和不平背景上的米粒图像说明了这一过程。这里使用了一个邻域，该邻域的 5×5 八边形中有 21 个像素，邻域中心与图像中的每个像素对齐。该区域中最暗的像素值取代原始的中心像素。通过使用原始图像的像素而非对其他像素应用该过程产生的新值，对图像中的每个像素重复执行这一操作。完成这一操作后，减小了米粒尺寸，如图 4.50(b)所示。重复该操作会继续减小米粒尺寸，并扩展基于局部背景亮度的背景。

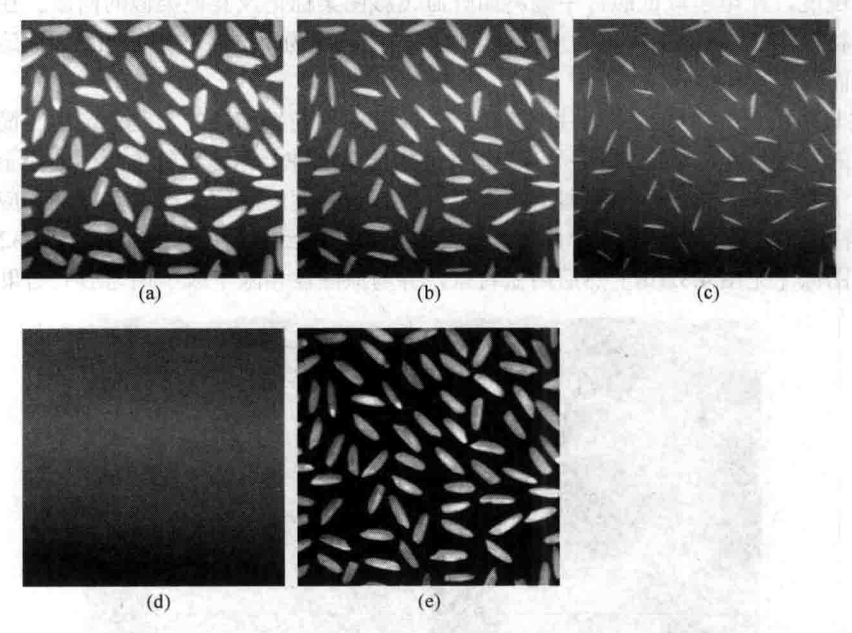


图 4.50 使用排序运算构造背景图像：(a)非均匀光照米粒图像；(b)用八边形 5×5 邻域中最暗邻域像素代替每个像素后的结果；(c)最暗邻域灰度膨胀运算的另一次重复结果；(d)4 次重复后仅保留了背景；(e)从图(a)中减去图(d)的结果

重复 4 次后［见图 4.50(d)］，完全移除了米粒。这种移除是可行的，因为任何米粒的最大宽度不会大于用于排序的 5 像素宽邻域半径的 4 倍。应用该方法需要知道所存在最大特征的宽度（最小尺寸），它必须小于所用邻域的半径，或者如果反复使用一个小邻域，可通过观察运算和重复的进展，直至完全移

除特征。这种方法产生的背景具有原始图像中存在的大规模变化，从原始图像中减去背景会产生调匀后的图像［见图 4.50(e)］，调匀后的图像清楚了显示了特征，并允许通过阈值处理方法将它们从背景中分离出来。在这个特例中，仅用到了最暗的值。通常会结合使用腐蚀和膨胀，如下面的几个例子所示。

基于排序的方法特别适合于在许多自然情形下出现的不规则背景亮度变化。图 4.51 显示了一幅月球的望远镜图像。表面与太阳夹角所致不均匀光照是引起亮度变化的原因之一，表面反射率的不同与高程的变化，也是引起亮度变化的原因。此时，应用灰度形态学“闭”运算（即一系列先膨胀再腐蚀操作），计算出了背景。第一步保留了 7 像素宽八边形邻域内最暗的像素（膨胀），接着保留了相同大小区域中的最亮值（腐蚀）。从原始图像中减去这个背景后，由于消除了整个变化，因而表面上各处的标记更为清晰。闭运算或开运算（先腐蚀再膨胀）的应用（取决于背景是否局部亮于或暗于感兴趣的特征），是所用的正确过程，而非一次腐蚀或膨胀本身，因为结构的整体尺寸得到了保留。

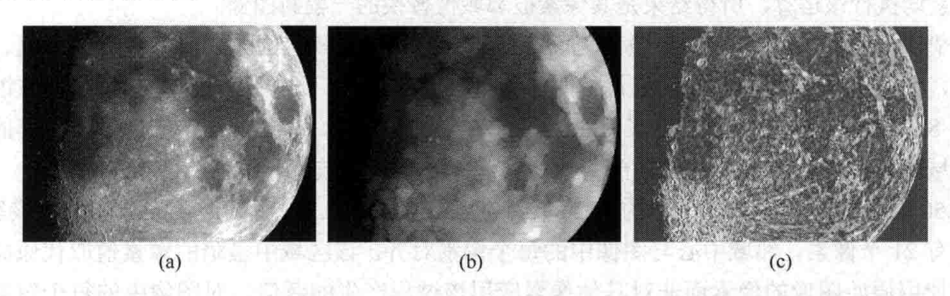


图 4.51　排序调匀应用：(a)月亮图像；(b)灰度闭运算生成的背景；(c)从原始图像中减去背景后的调匀结果

这种方法也有助于检查生物标本中冷冻断裂细胞壁上的粒子和其他表面粒子，检查不规则颗粒或花粉的表面粗糙度，在电子取证照片中隔离如鲜血飞溅图案标记及其他类似的问题。在某些情况下，它可用于增强材料 TEM 图像中失位的可见性，这些失位表现为不同谷粒中的暗线，而这些谷粒的整体亮度会因晶向而发生变化。

该技术尤其适用于标本如组织切片的穿透成像，其中不同细胞器的密度会产生于简单或平缓模式中不发生变化的背景强度，如图 4.52 所示。感兴趣物体是这种金标记实验中所用的小金粒。位于细胞质上的那些粒子要远亮于位于暗细胞器上的粒子，且基于亮度的阈值处理并不成功。应用灰度形态学开运算（对暗特征进行腐蚀随后，采用膨胀运算将细胞器恢复到其原始大小），如图 4.32 所示，消除了暗色金粒。该图像［见图 4.32(b)］然后用做背景，并与原图像相除生成了调匀后的结果。

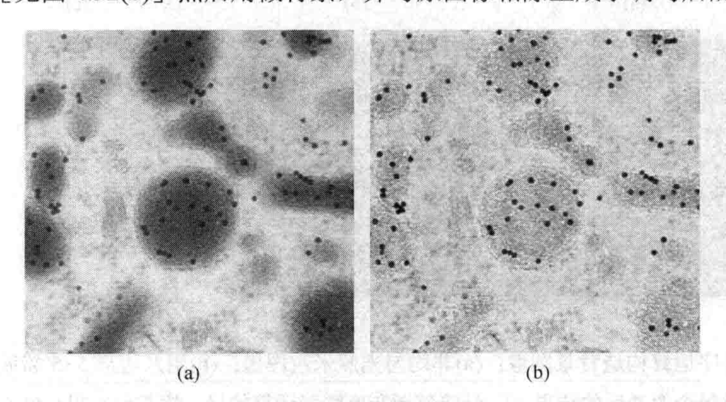


图4.52　穿透图像的调匀：(a)带有金粒的原始TEM图像［源自图4.32(a)］；(b)原始图像除以图4.32(b)所示背景图像后的调匀结果

图 4.53 给出了使用两种互补调匀技术的一个例子。首先，使用局部更亮的背景替换暗色轴突，生成一个排序后的背景。然后使用多项式拟合最亮和最暗像素值来局部扩展对比度，这有助于显示那些很细的暗色轴突线。

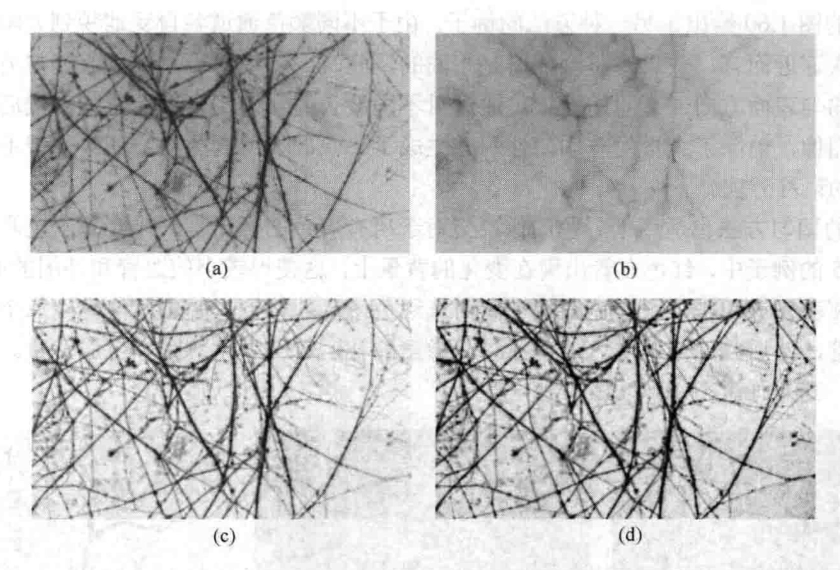


图 4.53　组合使用排序和多项式方法：(a)轴突的原始 TEM 图像；(b)使用 5 像素半径内最亮邻域像素替代暗像素后生成的背景；(c)图(a)除以图(b)的结果；(d)自动拟合多项式范围间的对比度扩展

减去（测量得到的、数学拟合得到的或图像处理得到的）背景图像来调匀亮度变化的能力，并不是一种无代价的过程。相减操作会用尽图像的部分动态范围或灰度。图 4.54 给出了一个例子。原始图像有一个阴影变化，这一变化用一个二次函数可以很好地拟合，但其动态范围约为 256 级灰度的一半。在减去函数后，调匀后的图像没有足够的剩余亮度范围去覆盖某些特征的暗色区域。这种“剪切”操作会影响到图像的进一步分析。在处理可能会用尽部分动态范围时，建议使用具有 8 比特亮度值的高动态范围图像。

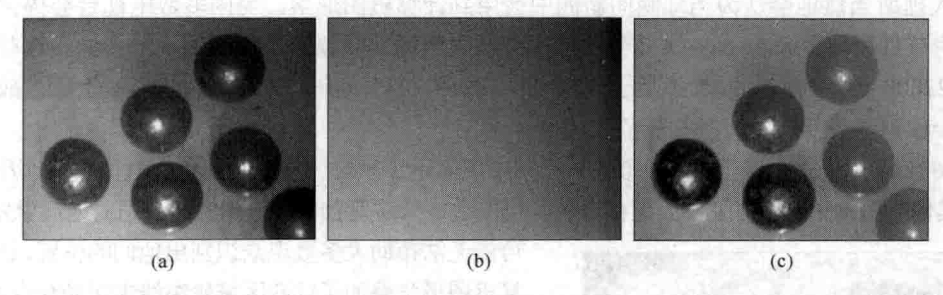


图 4.54　有限灰度范围图像调匀的效果：(a)原始图像；(b)拟合的多项式背景；(c)从图(a)中删除背景图像后的结果。背景是均匀的，但暗色特征不均匀，因为原始图像中的原始像素完全是黑色的

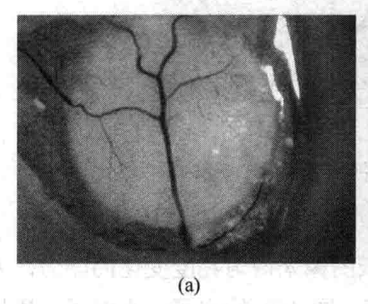
4.10　彩色图像

对于阴影校正，彩色图像有时会出现明显的问题。典型的例子出现在航空或卫星图像中，此时地面的不规则或星球表面的曲率会产生一种整体阴影。有时，这只影响图像的亮度而不影响颜色信息。但是，取决于相机的响应、可能的纹理、来自表面的镜面反射、大气吸收以及成像的其他细节，图像中不同区域间也会出现色偏。

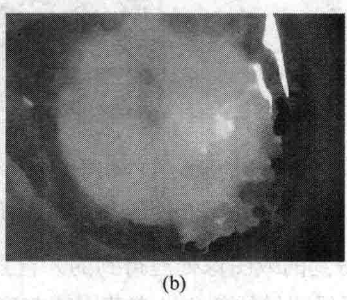
上面用于灰度图像方法，同样适用于彩色图像的亮度通道。调匀操作实际上不适用于红色、绿色和蓝色通道，因为它会产生改变图像的色偏。

第 1 章中的图 1.60 给出了另一种方法的例子。由于不同颜色通道各自受地貌到太阳的局部方向影响，因此对一次逼近而言，效果对整个色谱是相同的。此时，应使用一个颜色通道与另一个颜色通道之比来调匀不均匀表面方向或光照的效果。通过对不同波长的彩色图像进行滤波，然后用一幅图像去除以另一个幅图像，消除了大量的非均匀性，并生成了一幅调匀后的图像，其中位于不同区域的类似特征具有相同的最终外观。

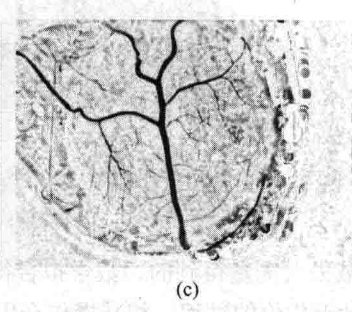
基于排序的调匀方法也适用于彩色图像，但对于所有彩色图像而言，排序基于色彩空间中的矢量距离。在图 4.55 的例子中，红色血管出现在变化的背景上，这使得较小的血管和外围的血管难以检测。先使用来自其邻域（该邻域位于彩色空间中离所选红色最远的位置）的彩色值替换每个像素（颜色腐蚀）来删除血管，然后使用最近颜色值替换每个像素的补运算（颜色膨胀）生成背景。从原始图像中减去背景图像，变为血管产生了较好的对比度。



(a)



(b)



(c)

图 4.55　文中描述彩色图像的排序调匀示例：(a)原始图像；(b)颜色开运算生成的背景；(c)图(a)减去图(b)后的结果

4.11　非平面视图

多数人理所当然地会认为当地晚间新闻中会用到计算机图形学。美国多数电视台会提供天气预报节目，这些节目利用了来自 NOAA（美国国家海洋大气局）气象卫星发回的卫星图像。这些图片展示了高度为 22000 英里的地球同步卫星于北纬 0°、西经 108°（有时为获得南大西洋上所生成飓风的更好视图，会移到西经 98°）时所拍摄的美国。

这些图像显示了云图，且白天拍摄的一系列图像显示了风暴的移动和其他天气变化。海岸线、北美五大湖和其他一些地形特征很明显，但部分特征可能会被云层遮挡。考虑到普通公民的地理学知识，图片可能无法帮助大多数观众识别出他们的位置，因此使用计算机图形学叠加了行政区域轮廓线来帮助观众进行识别，譬如州界和其他信息，如市、县或主要的高速公路。多数美国电视台在广告、新闻领域的计算机图形学方面不惜重金投入，但很少自己生成这些行政区划线，而只是通过专业市场中的某家公司来得到已有行政区划线的图像。

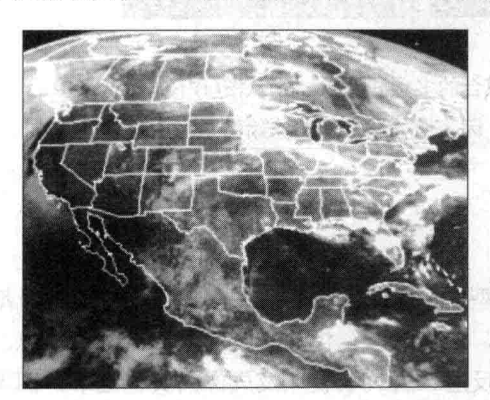


图 4.56　叠加有行政区划线的北美 GOES-7 图像。加州 Baja 西部的暗色区域是在 1991 年 6 月 11 日日食期间月亮的阴影

生成这种行政区划线并不只是简单地将传统地图（如学校教室悬挂的墨卡托投影地图）叠加到卫星图像上，而是需要考虑地球的曲率和图像的透视效果。图 4.56 显示了一幅北美地区的气象卫星图像，该图像顶部缩短，并显示了从西到东的明显曲率。地面上一个正方形看上去如同一个倾斜的梯形，而且如果正方形足够大，则其两条侧边会有明显的弯曲。

地面上点的经纬度坐标用于计算球体的透视图，如同从卫星上看到的那样。由于观察点是固定的，因此这是一次性计算，但为叠置目录构建较好的轮廓图，则需要对大量的点进行这一计算。图 4.57 中以图形方式显示了这一计算。

地球上某点的位置（由该点的经纬度指定）用来求出至卫星的视线与球体前方插入的图像平面的交点。这一计算只需简单的三角测量学，如图 4.58 所示。该平面上点的坐标是被观察图像中该点的位置。

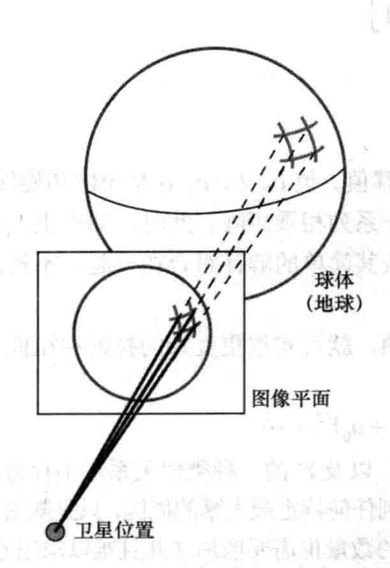


图 4.57 卫星成像图解。如同任何透视几何一样，"平面"图像形成于从三维物体的观察线投影到观察点，并从它们与图像平面相交的点来构建图像

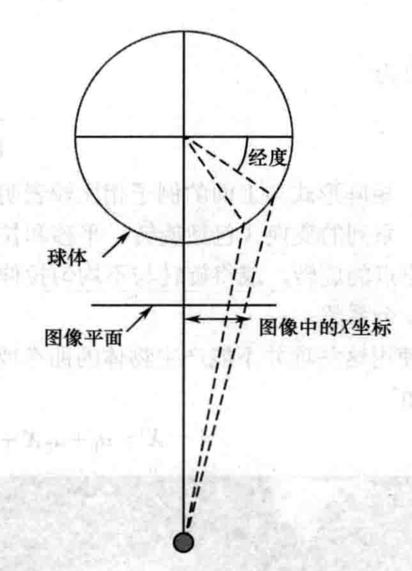


图 4.58 由地面上某点的经度和卫星位置来计算图像平面中该点位置的三角测量法。这是来自北极的视图；来自赤道的类似视图会由纬度给出 y 坐标

4.12 计算机图形学

计算机图形学常用来构建透视图和三维物体渲染，以便可在计算机屏幕上观看，如在 CAD（计算机辅助设计）程序中。对此主题的讨论超出了本书所需，感兴趣的读者可参阅如 Foley & Van Dam（1984, 1995）或 Hearn & Baker（1986）这样的标准教材。显示处理与前述是一致的，只是还加上了透视控制，透视控制允许用户调整相机或观察点的距离，进而控制透视收缩的程度（相当于为相机选择一个长或短焦镜头，短焦镜头在图像中产生更多的失真）。

忽略透视收缩失真（如使用一个长焦镜头），三维情形下一个点的平移可以用其 x、y、z 坐标与对应于旋转和平移的一组值的矩阵相乘来表示。在二维情形下这更为简单，因为这里的主要目的是处理二维图像。考虑笛卡儿坐标为(X, Y)的点以及这些点的移动或旋转方式。

物体平移会对 X 和 Y 添加偏移量，产生

$$\begin{aligned} X' &= X + \Delta X \\ Y' &= Y + \Delta Y \end{aligned} \tag{4.10}$$

而拉伸物体可能需要不同的乘法系数：

$$\begin{aligned} X' &= \alpha X \\ Y' &= \beta Y \end{aligned} \tag{4.11}$$

将物体旋转 θ 角会在原 X 和 Y 坐标间引入如下形式的依赖关系：

$$X' = X\cos\theta - Y\sin\theta \qquad (4.12)$$
$$Y' = X\sin\theta + Y\cos\theta$$

总之，二维平移表示最常写为齐次坐标和矩阵形式。坐标 X, Y 连同一个任意常数 1 被组合到一个矢量中，进而允许平移值被包含到矩阵数学中，产生结果

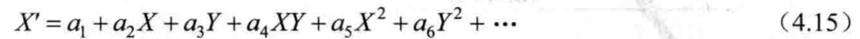
$$[X' \ Y' \ 1] = [X \ Y \ 1]\cdot\begin{bmatrix} a & b & 0 \\ c & d & 0 \\ e & f & 1 \end{bmatrix} \qquad (4.13)$$

相乘结果为

$$X' = aX + cY + e \qquad (4.14)$$
$$Y' = bX + dY + f$$

这一矩阵形式与上面的例子相比较表明，e 项和 f 项是平移值。值 a、b、c、d 是拉伸和旋转系数。当组合一系列的变换（包括旋转、平移和拉伸）时，会产生一系列相乘矩阵。此时，如产生关于某个点而非原点的旋转，或将旋转与不均匀拉伸相组合，各项会按其简单的解释组合在一起。不管怎样，只需要 6 个系数。

仅使用这些项并不能产生物体的曲率或扭曲。引入高阶项，就可实现更复杂的拉伸和扭曲。这会产生形如



$$X' = a_1 + a_2X + a_3Y + a_4XY + a_5X^2 + a_6Y^2 + \cdots \qquad (4.15)$$

的复杂方程式，以及 Y' 的一种类似关系。不存将这个多项式展开限制到任何特定最大幂的原因，只是随着复杂性的增加，系数的数量也需要增加（并且难以求出它们），且应用变换所需的数学精度也会增长。二次幂以上的项并不常见，二次幂就足以处理多数常见的失真，甚至可以在小角度范围内逼近从卫星上观看球体表面时产生的曲率。

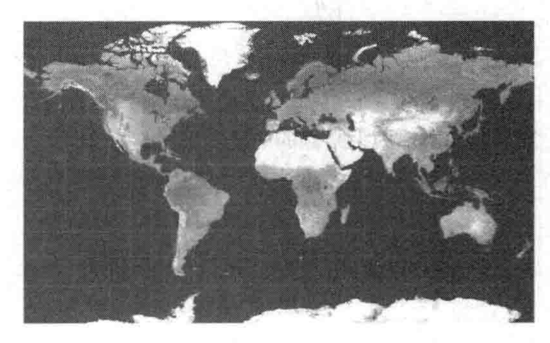


图 4.59　地图投影所用的标准墨卡托投影，此时地球上的点投影到圆柱体上，在高纬度地区产生了明显的失真

当然，有些表面映射可由其他函数更好地处理。球体地球在圆柱上的标准墨卡托投影（见图 4.59）会将两极设置为无穷远处，并会放大高纬度地区。为逼近这一投影，需要许多多项式项，但由于该映射的实际几何形状是已知的，因此可有效并正确地使用余割函数来实现这一变换。

4.13　几何失真

本书中的这些数学运算对于将图像转换到可进行精确测量而言，是有意义的。所得到的图像从垂直视角看起来通常并不是平坦的表面。很明显，上文所用的卫星图像就是这样的一种情况。很多犯罪现场照片和监控图像需要从非垂直视角拍摄。在电子显微镜中观看标本时，通常会倾斜标本表面以增强图像对比度。如果表面不平，不同区域可能会以任意角度倾斜，或可能会出现连续的曲率。由于移动或侧扫模式的使用，或将单条测线成像到连续移动的胶片上，许多机载相机和雷达会引入可预测的失真（因而可纠正）。在所有这些情形下，知道失真是校正它的关键。

这种情形对于光学显微镜并不常见，因为光学景深很浅，故表面必须是平的且与光轴垂直来保持对焦。但也存在其他成像技术，它们适用于经常遇到的非理想表面或观看条件。

航拍或卫星图像成图就是这样一种应用（Thompson，1966）。没有一种完美的投影能将球体表面

展开到平面上，并保持距离和方向不变，因此会用到各种近似。在每种情形下，球体坐标和地图坐标间都存在一种可用数学方式表达的已知关系。如果确切知道观察点（如气象卫星这一情形），或曝光时可以计算出观察点（如探测器经过某颗行星这一情形），则可以求出这一数学关系。

由于飞机位置并不能精确地控制，虽然图像拍摄的位置可以通过 GPS 记录，但生成的校正图像对于航拍照片而言通常并不实用。另一种选择是在图像中定位几个参考点，这些参考点的位置在地球或地图上是已知的，因此可使用它们来求出图像上位置和地图上位置的关联方程。这种技术通常被称为图像扭曲或橡皮拉伸，而尽管关联方程与计算机图形学中所用的那些方程一样，但用来求出系数的技术是不同的。

由原始坐标 X, Y 计算变换视图坐标 X', Y' 的一对方程，可包含常量、X 和 Y 的线性项，以及诸如 XY、X^2 等之类的高阶项。添加更多的高阶项会在变换中引入更多的复杂失真。若问题只是旋转，则只需要线性项，且为保角目的可引入关于系统的约束条件。按照式（4.11），这就要求拉伸系数必须相等。因此，只需要少数几个常量，并且它们可以通过定位几个已知的参考点及建立联立方程来求取。

图像彼此配准这种更为精确的拉伸，相应地要求更多的项和更多的点。在电子显微镜中，大景深允许获取局部平坦但与视点成某个角度的标本照片，但本质上产生了梯形失真，如图 4.60 所示。最靠近镜头表面的部分与更远的区域相比，放大得更多，而距离在倾斜方向上则被透视收缩。为了测量和比较这些表面上的特征，或甚至正确地应用图像处理方法（通常假设不同方向的相邻像素与中心像素等距），可能需要对这幅图像进行变换来校正失真。由于确切的倾斜角和工作距离可能未知，因此需要一种仅使用图像自身内的参考点的方法。

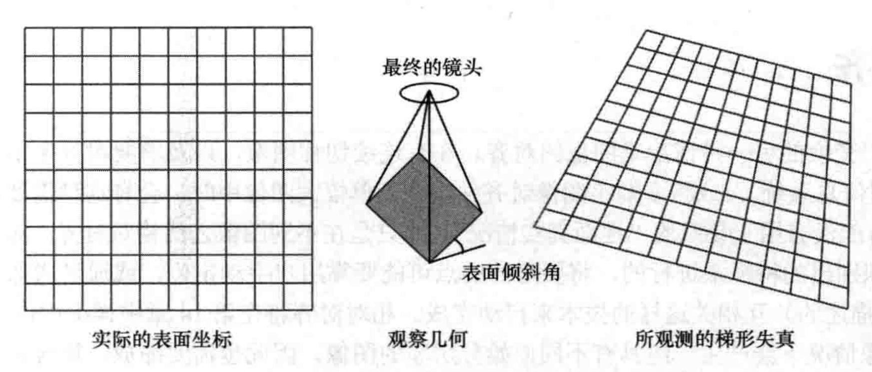


图 4.60　用短焦距观察倾斜表面时经常遇到的梯形失真

所需要的是识别表面上其实际 X, Y 坐标已知且其图像坐标 X'，Y' 可以测量的 4 个点。然后对 4 组坐标中的每一个，可写出下面的方程组：

$$X = a_1 + a_2 X' + a_3 Y' + a_4 X'Y' \quad (4.16)$$
$$Y = b_1 + b_2 X' + b_3 Y' + b_4 X'Y'$$

这允许求解出常数 a_i 和 b_i。当然，如果有更多的点可用，则可用来求出最小平方解，进而使得图像中坐标测量小误差的影响最小化。无论是期望的还是观察到的失真，都在模型方程中按照计算效率（所需参考点的数目）和系数的精度使用尽可能少的项。

图 4.61 所示例子中，由犯罪现场拍摄的轮胎胎面标记照片可以看出胎面以一个角度压印。由该视图足以辨别轮胎的制造商和型号，但不便于测量用来与特定轮胎胎面压印匹配的轻微磨损标记的位置。在图像中标记一个其边缘平行于轮胎压印的区域，就定义了观察点和失真，因此该视图可按所示方式进行校正。

对特征来说，进行透视校正时出现的主要问题并不在倾斜的平面上，而是在其前或后面延伸。在图 4.62 的例子中，建筑物的正面已被校正，以显示结构的正确角度和尺寸，但前面街道对面的灯柱看起来与建筑物有一个角度，而灯柱实际上是与建筑物垂直的。还要注意的是，建筑物右侧看上去聚焦不清晰，这是由于该部分图像已被放大和内插。内插技术将在下面讨论。

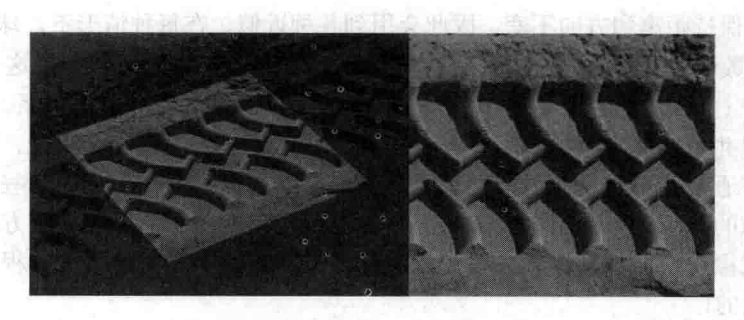
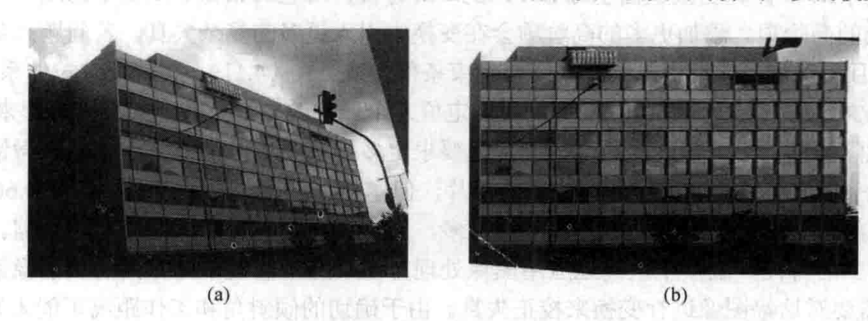


图 4.61　画出平行于轮胎轨迹边缘的轮廓定义了图像失真，由其可计算出必要的校正



(a)　　(b)

图 4.62　校正由非垂直观察所产生的梯形失真

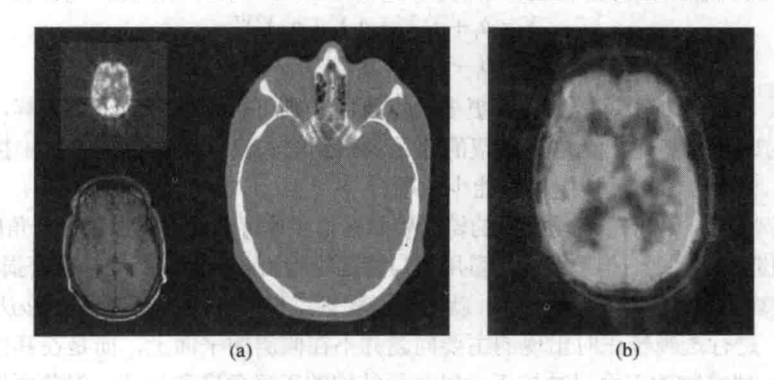
4.14　对齐

需要几何变换的另一种情形是图像的对齐。在有连续切片图像，以及需要对齐多幅表面地图而建立 GIS（地理信息系统），或将多幅小图像对齐并组合到单幅大图像中时，会有这种需要。在某些情况下，对齐时可能没有用于参照的“地面真实情况”，而只是在不同图像之间相对对齐。对齐是通过使用图像中可以识别出的特征来进行的。将要对齐的点可能要靠用户手动定位，或通过成像系统采用诸如（在第 6 章中描述的）互相关这样的技术来自动完成。相对对齐将在第 14 章中详细讨论。

有些成像情况下会产生一组具有不同原始分辨率的图像，因此也需要缩放、旋转和对齐。此时需要具有三个常数（和 Y 的一个相似方程式）的对齐方程：

$$X = a_1 + a_2 X' + a_3 Y' \tag{4.17}$$

图 4.63 显示了一组图像，各图像中的三个点已标记为对齐参考点，图像的变换为拉伸和旋转。



(a)　　(b)

图 4.63　不同尺度的三幅图像的配准与对齐：(a)PET（小）、MRI（中）和 CT（大）头部扫描头像。红色、绿色和蓝色点是配准点，用于确定拉伸和旋转矢量；(b)配准后的结果，图像显示在各自的彩色通道中

由于刀片所致的块压缩，在显微镜下用切片机从嵌入的生物体切下的连续切片，一般会在切削方向缩短 5%～15%。在这些切片成像前，通常会任意旋转切片。可以在标本中引入基准标记（如穿过标本块的孔或切割截面前插入的纤维）来帮助定位匹配点。

沿图像边缘对齐叠加图像可能需要大量的变形，进而将它们拼接到一起（Milgram，1975）。逐边切片对齐的拼接，通常会使用广角图像（包括在电子显微镜下的典型情况），但会因为上述梯形失真而失败。所得结果是，沿图像边界的特征并未完全对齐，因而拼接并不完美。使用橡皮拉伸可以校正这一缺陷。

这类校正通常用于卫星和空间探测器图片。图 4.64 显示了拼接由勇气号火星探测器在火星上所拍多幅图像的例子。近年来，人们常由单独的图像来创键拼接，进而生成全景视图，以扩大数码相机的分辨率。图 4.65 显示了在某个位置（通过旋转三角架上的相机）所拍摄的 4 幅图像，它们形成了一个有较大重叠的 2×2 阵列。广角镜头在每幅图像中产生了失真，而且视野的简单叠加并不能对齐图像。图 4.66 显示了最终的拼接，即建筑物的一幅完整图像。注意，为平滑地匹配在一起，图像已变形。这类拼接对于自动创建全景照片是非常有用的（Brown & Lowe，2003；Brown et al.，2005；Chow et al.，2006；Kim & Hong，2006），但对于测量目的基本无用。

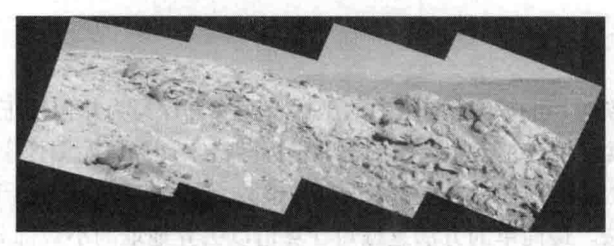


图 4.64　火星上古谢夫陨坑的拼接图像，各幅图像由勇气号火星探测器拍摄



图 4.65　叠加 4 幅图像提供了覆盖更大场景的 2×2 阵列。注意在每幅图像中的不同角度和不兼容镜头失真

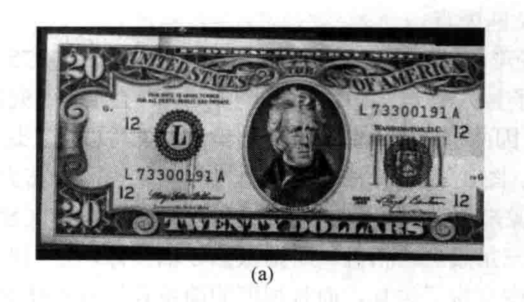


图 4.66　图 4.65 中 4 幅图像产生的拼接效果

对于科学成像目的，使用显微镜载片台或其他标本定位装置，可以以合理的精度移动被成像的标本而相机保持固定，进而可以获取无限大小的图像（Bacus 1995，2000，2001；Gu & Ogilvie，2005；Thevenaz & Unser，2006）。在这种情况下，可用约束条件帮助将图像匹配在一起。例如，已知图像可轻微旋转，但却并未失真（即直线仍为直线，且角度不变）。因此，组合在一起的过程只是整体上移动和旋转各幅小图像。

如果小图像间重叠 10%～20%且角度失配不超过几度，那么将每幅小图像匹配在一起会产生尺寸较大的高分辨率拼接，如图 4.67 所示。匹配技术基于互相关（见第 6 章），为实现最佳匹配，使用了一种迭代过程来将所有小图像拼接在一起。对使用原子力显微镜获得的图像，这尤其有效，因为这些设

备所覆盖的区域通常很小，且由于具有很高的空间分辨率，设计绝对精确移动标本的硬件并不可行（Condeco et al.，2000）。



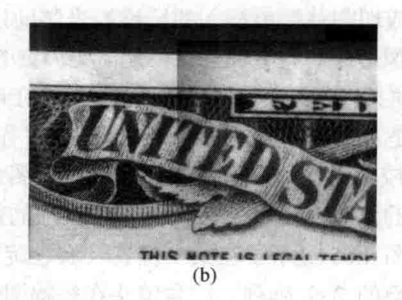


图 4.67　(a)由 8 幅小图像拼接形成的大图像，数字相机所摄每幅小图像的尺寸为 3200×2400 像素；(b)两幅小图像匹配的细节

4.15　内插

对图像进行几何变换时，可按原始图像中的坐标将变换方程写为几何校正后图像中的坐标的函数，反之亦然。实际中，通常使用校正后图像中的 x, y 坐标网格来计算原始图像中每个像素的相应坐标。

遗憾的是，为原始位置计算出的这些坐标很少为整数。这意味着位置位于原始图像的像素“之间”。处理这一问题有几种方法。最简单的方法是截短计算值以丢弃地址的小数部分，而采用接近坐标系原点的像素。将地址值舍入为最近的像素，然后将该像素的亮度复制到变换图像阵列中，可得到稍好一些的结果。这就是所谓的“最近邻”过程。如图 4.68 所示，这种方法在测试图案中产生了混叠或“阶梯状”线条，且宽度有明显的变化。

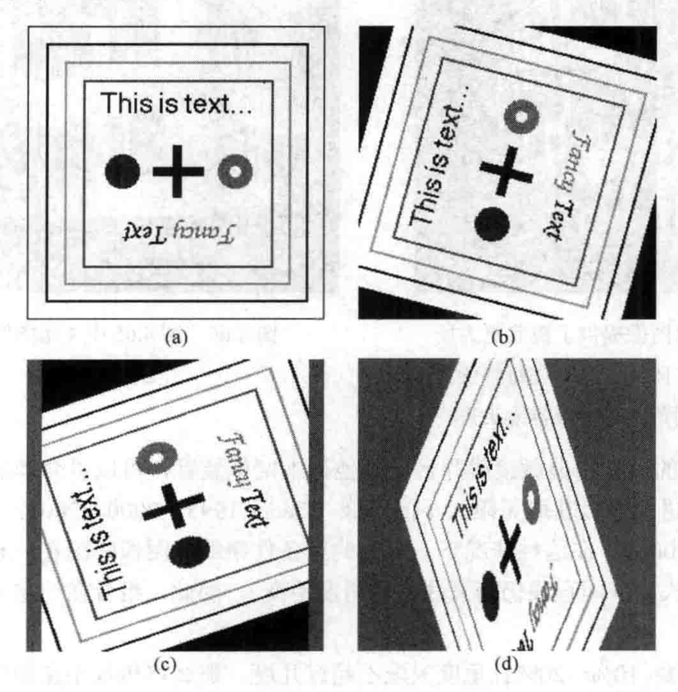


图 4.68　使用最近邻像素值法旋转并拉伸测试图像：(a)原始图像；(b)仅旋转但尺寸无变化的结果；(c)旋转和均匀拉伸但保角的结果；(d)旋转和均匀拉伸但不保角的结果（线条仍为直线）

当这种失真不可接受时，可使用内插。变换后像素的亮度值可通过在围绕所算地址四周的 4 个像素间进行内插来计算。这被称为双线性插值，它通过 X 和 Y 坐标的小数部分计算。内插首先在一个方向上进行，然后在另一个方向上进行，如图 4.69 所示。对于坐标为 $j+x$，$k+y$ 的位置（其中 x 和 y 是地址的小数部分），第一次内插的方程为

$$B_{j+x,k} = (1-x)\cdot B_{j,k} + x\cdot B_{j+1,k} \tag{4.18}$$
$$B_{j+x,k+1} = (1-x)\cdot B_{j,k+1} + x\cdot B_{j+1,k+1}$$

然后 y 方向上的第二次内插值给出最终值为

$$B_{j+x,k+y} = (1-y)\cdot B_{j+x,k} + y\cdot B_{j+1,k+1} \tag{4.19}$$

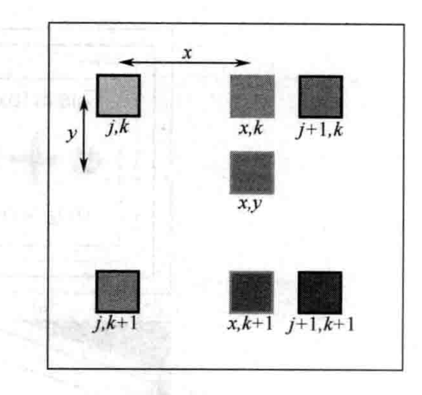


图 4.69　双线性像素内插图解。使用像素地址的小数部分，邻像素的亮度值首先水平地内插以求出以红色标记的位置的亮度值，然后这两个值进行垂直内插，以求出以蓝色标记的目标像素的亮度

有些情形下，也会对较大的区域使用加权内插，使用最普遍的是双三次拟合。双线性内插采用相邻像素值的 2×2 阵列来计算内插后的值，但双三次拟合方法采用 4×4 阵列。通过使用与式（4.18）和式（4.19）给出的相同表示，从 $k-1$ 到 $k+2$ 及从 $j-1$ 到 $j+2$ 求和。水平内插的中间值是

$$B_{j+x,k} = (1/6)(B_{j-1,k}R_1 + B_{j,k}R_2 + B_{j+1,k}R_3 + B_{j+2,k}R_4) \tag{4.20}$$

垂直方向的内插值为

$$B_{j+x,k+y} = (1/6)(B_{j+x,k-1}R_1 + B_{j+x,k}R_2 + B_{j+x,k+1}R_3 + B_{j+x,k+2}R_4) \tag{4.21}$$

其中权重因子 R_i 由地址的实部（x 或 y 分别进行）计算为

$$\begin{aligned} R_1 &= (3+x)^3 - 4(2+x)^3 + 6(1+x)^3 - 4x^3 \\ R_2 &= (2+x)^3 - 4\,(1+x)^3 + 6\,x^3 \\ R_3 &= (1+x)^3 - 4x^3 \\ R_4 &= x^3 \end{aligned} \tag{4.22}$$

双三次拟合与双线性方法相比，各向同性更好，因而可使得导数在被拟合像素附近的 4×4 区域的角点处连续。内插总是会平滑图像并删除一些高频信息，但可消除沿直线或边缘的混叠或“阶梯状”效果。需要时，可以改变式（4.22）中的系数，以产生更清晰或更平滑的结果。也可使用更大的邻域，但内插通常是在傅里叶空间中执行，而非直接访问邻近像素。

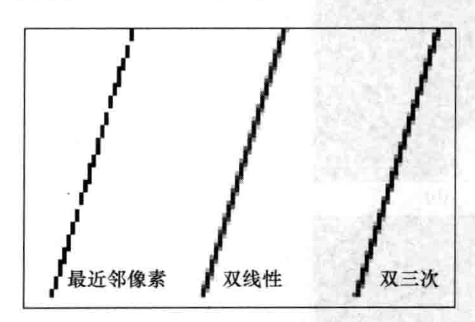


图 4.70　采用最近邻像素、双线性和双三次内插旋转 1 像素宽的黑线的效果比较

图 4.70 比较了无内插（选择最邻近的像素值）、双线性内插、双三次内插将线条（最初为 1 像素宽的黑色垂直线）旋转 17° 后的结果。最近邻方法产生的混叠现象很明显。双线性内插结果与双三次内插结果相比，对比度下降更为明显，但双三次内插对线宽的模糊更甚，这两种方法都将灰度值赋给相邻像来平滑线条的外观。

内插的优点是尺寸在变换中改变得很少，且边界和其他线条不偏离或失真。图 4.71 显示了与图 4.68 相同的例子，只是使用的是双线性内插。由图可知，线条看起来笔直且未出现混叠或“阶梯状”现象，因为沿线条两侧的一些像素具有由内插值生成的中间灰度值。计算机图形学程序通常使用这种方法在 CRT 显示屏上显示线条和字符，因此避免了离散像素阵列所固有的阶梯状现象。这种技术称为反混叠，有时由显示硬件处理，并根据它们与线条数学位置的远近程度来生成带有灰度值的像素，进而使得观看者看到平滑线条或字符的错觉。

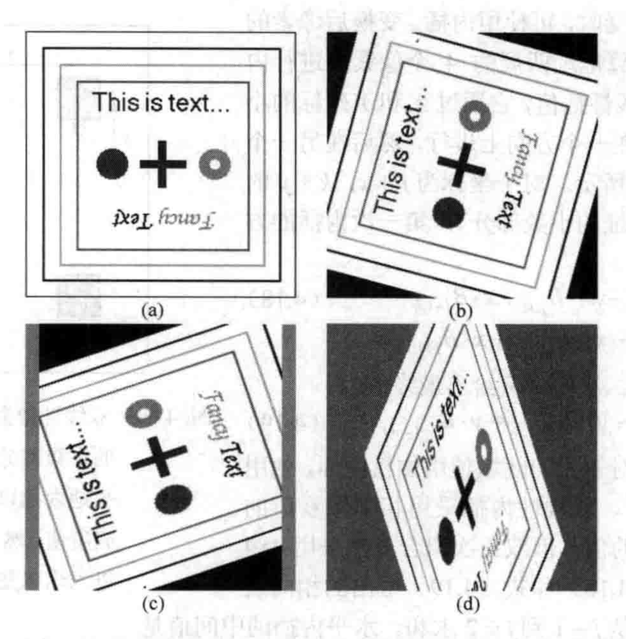


图 4.71　与图 4.68 中相同的旋转和拉伸，但采用的是双线性内插法。注意线条和边界的平滑度

也可以使用高阶多项式或自适应样条拟合像素亮度值。放大图像时，为降低由于锐利边缘被传统内插展开所引起的感知模糊，这可能特别有用。图 4.72 显示了一个例子，其中分别用无内插、双线性内插、自适应样条拟合和分形内插实现了 4 倍放大。后者往图像里插入了虚假的“细节”(Freeman et al.，2001；Fattal，2007；Chen et al.，2008)，而样条拟合则为边缘保留了最佳的视觉清晰度。

对于图像扭曲、对齐或放大，内插具有保形的优点，但会使得亮度值变化。使用对像素地址舍入得到的最近邻像素法，结果图像的尺寸会失真，但亮度值不变。为特定成像任务选择哪种合适方法，主要取决于哪类信息更为重要，其次取决于内插所需的额外计算代价。

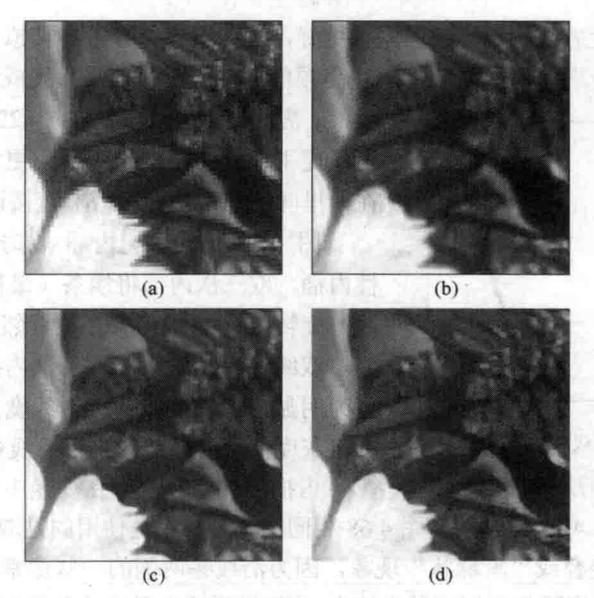


图 4.72　放大一幅图像（所显示部分与原图相比放大了 4 倍）：(a)无内插（最近邻）；(b)双线性内插；(c)样条拟合；(d)分形内插

图 4.73 给出了对扭曲方程添加高次项后的效果。采用二次项，可校正短焦镜头或 SEM 的梯形失真。也可在适度的距离上对球面的失真进行建模。使用高阶项，任意失真都是可能的，但因为不太可能有多个参考点来确定这样一种失真，因此在图像处理中很少使用。

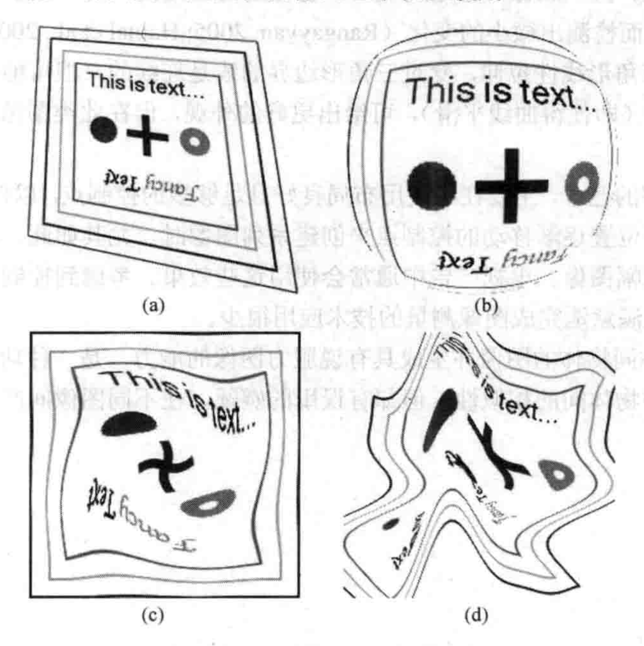


图 4.73 使用与图 4.68 相同原始图像的另外一些图像扭曲示例：(a)具有梯形透视缩短的二次扭曲；(b)直线变弯的二次扭曲（近似为一个球面）；(c)视野中心但保角的扭曲（也是三次扭曲）；(d)需要高阶项和三角函数项的任何扭曲

4.16 变形

可根据数学关系实现可控扭曲的程序，或从应用到整幅图像上的一组识别基准或参考标记来计算矩阵值的程序，通常相当专业。目前，已有基于一组用户定义的控制点来实现图像变形的程序。控制点通常置于两幅图像中相应的特殊位置。例如，为对齐两副面孔，可使用位于眼睛末端、嘴角、沿发际线和颌角线的点作为控制点，如图 4.74 所示。

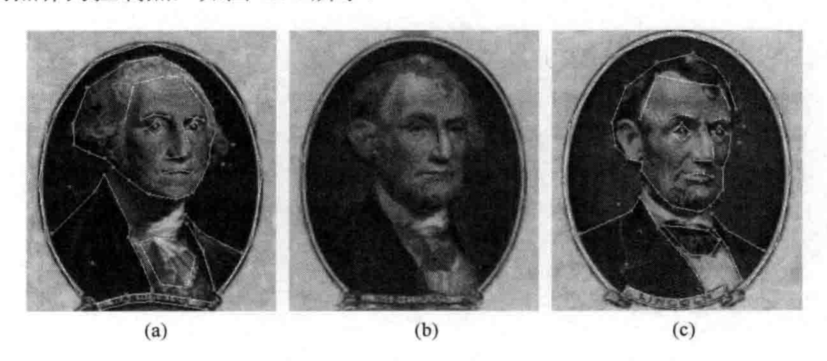


图 4.74 华盛顿至林肯的变换。每幅原始图像上标记的相应点控制从一幅图像到另一幅图像的渐进变形。中间帧似乎为一个人，它具有左右两帧的特征。注意图像下部的横条，在两幅原始图像中其的弯曲方向不同，但在合成图像中却基本上是直的。实际应用中，会使用更多的对应点来产生更好的结果

该程序使用这些控制点来将第一幅图像中的曲面分解为三角形，具体选择哪些点作为每个三角形的角点则由称为 Voronoi 图（也称为泰森多边形）的程序确定。每个三角形一致地拉伸，以便与第二幅图像中的角点位置匹配。医学成像常用这类变形，它使用内部结构作为一组参考点，目的是将患者图像与参考标准对齐，进而检测出微小的变化（Rangayyan, 2005; Hajnal et al., 2001）。

对三角网和每个三角形线性拉伸，穿过三角形边界的线是连续的，但可能会急剧弯曲。使用样条或三次方程来控制拉伸（即使得曲线平滑），可给出更好的外观，但在此类图像上进行测量通常要牺牲一些保形效果。

使用这些变形程序的艺术，主要在于使用布局良好且足够多的控制点，以使结果看起来更为逼真。使用从原始位置到最终位置逐渐移动的控制点来创建系列图像时，尤其如此。这些变形的“影片”会将一幅图像渐变到第二幅图像。电视广告中通常会使用这些效果。考虑到控制点设置的任意性及尺寸和方向的失真，能令人满意地完成图像测量的技术应用很少。

使用变形来对齐不同物体的图像并生成具有说服力图像的能力，是一种功能强大的演示工具，这种工具可用于显示两个物体间的相似性，但却有误用的嫌疑，在不同图像间产生明显匹配实际上并不相同。

第5章 空间域图像增强

前面的章节讨论了如何校正或减轻图像获取时的主要缺陷。校正这些缺陷与进一步增强图像之间仍有个模糊地带。本章的主题是增强。可用方法是增加图像某部分、某方面或某成分的可见性，这通常是靠抑制其他部分可见性来实现的。因此，图像处理有类似于文字处理或食品加工。它可能会重新排列事物，使产品更加令人满意或能够解释，但总数据量并不改变。

至于图像，这通常意味着字节（或像素）数量并不减少，但个体值可能会改变或重新排列。与图像处理相比，后续章节主题的多数图像分析程序试图仅从图像中提取"重要的"信息。例如，鉴别和计算某幅图像上的对象时，会将数据量从大约100万字节减少到几十字节，或者甚至在某些质量控制、医学或取证应用中，只是得出简单的"是"或"否"答案。

进行图像增强的原因有多种。原因之一是仅使得图像更容易形象化研究和解释。这里描述的许多程序一定程度上都基于人类视觉系统的响应和需求。有些程序则是被证明行之有效的方法。其他方法基于图像生成的物理原理（如光与物体的相互作用），或基于光学元件和图像检测器的操作（如消除镜头失真或校正摄像机的响应）。后者应用起来不一定更复杂。例如，Sharma（2005）给出了从扫描文档中消除"透印"的计算方法，该方法基于光通过文档的路径及光从两个表面的不同反射来区分图像。原因之二是使得图像打印更容易，或允许使用自动方法来进行测量，这些内容将在后面几章中介绍。

图像增强处理可以在空间域（由图像的常规表示组成的像素阵列）或其他域（如第6章中讨论的傅里叶域）中进行。在空间域中，像素值可以根据与原始像素值相关的规则（局部或点处理）来修改。另一类运算会以各种方式与其局部邻域中的其他像素值进行组合或比较。第4章中出现这些方法的例子，目的在于取代亮度值，进而拉伸图像的对比度，或通过核平均或中值滤波的方法来平滑噪声。本章所用的技术采用同样的基本工具集来进行进一步的增强。

值得注意的是，二维图像的像素个数通常很多（一般为几百万像素）。即使根据之前内容简单修改每个像素值的一个点运算，也需要计算机对每个像素位置进行寻址。对于邻域运算而言，由于必须多次寻址每个像素，因而处理速度会变慢。此类处理运算的基本要求是，要有快速处理器和高速内存存取（及大量的内存），目前多数现代计算机可以在很短的时间（通常不到1秒）内完成运算。性价比合适时，可使用某些专用机器（如移位寄存器和阵列处理器、具有专用存储器和定制寻址电路的电路板、具有特殊编程的多处理器）来实时处理图像。随着CPU速度的提升，这种特殊硬件已很少使用。使用更大的内存，可以迅速地从磁盘存储器上一次性存取整幅图像，而无须将文件分成许多小块来处理。

随着个人计算机功能的增强，出现了用于图像增强的复杂算法。此类算法中的许多并不是新开发的，而是源自几十年前，且这些算法随着卫星和空间探测器图像处理的需求而发展，但直到最近才可以用大型计算机或专用系统实现。多数此类算法在台式机或笔记本电脑中处理图像时，可在几秒钟内完成。有些算法甚至可在消费类数码相机的固件中实现。

本书并不关注加速各种处理运算的实现方法，但要记住，这种编码"技巧"在30年前甚至15年前，是使得计算机图像处理可实用化的重要方法。随着更快的计算机、更大的内存和更智能的编译器的出现，人们对这类需求的关注更少了。随着更加复杂算法的实现，通常而言，与生成尽管能提升少量性能但不能轻易于平台间移植的专用程序相比，程序员写出清晰且文档齐全的代码更为重要。程序员应有动力来精确地实现算法，但时至今日，程序设计在某种程度上仍未达到人们的期望。

5.1 对比度控制

第 4 章中给出了通过重新分配像素的亮度级来增强暗淡图像对比度的例子。在许多系统中，这可以通过将一个数值表写入到显示硬件中来完成。这个查找表或 LUT 对于每个存储值，用一个显示亮度值替换，因此不需要修改内存中存储的用于图像的任何值。通过令最暗的像素值为黑色，最亮的像素值为白色，并使用线性内插后的灰度来扩大对比度范围，很好地利用了显示屏并增强了图像中特征的可见性。

第 1 章表明，相同的 LUT 方法可以通过将红色、绿色和蓝色值的三元组指赋给每个存储的灰度值，而用于彩色图像。这种伪彩色也增加了相似像素之间的差异，有时它对希望看见或显示图像亮度中小变化或逐渐变化的用户有一定的帮助。

计算机显示屏通常可以显示 2^8 种或 256 种不同的灰度，且对于每个红色、绿色和蓝色分量，同样可以产生 2^8 种亮度值的色彩，百而形成总共 2^{24} 种或 1600 万种不同的色彩。因为用来显示的颜色足以表示绝大多数自然场景，因此常称这种彩色为“真彩色”。当然，这并不意味着所显示的颜色在光测量上是准确的，或者等于场景中的原始颜色，或者显示屏能够显示眼睛能够感知或相机能够捕捉的色彩的全色域。要做到上面这一点，事实上非常困难，并且需要特殊的硬件和校准。若原始图像在每个颜色通道上的亮度值比 256 种更多（超过 8 位深度），那么此时为了能在屏幕上显示它，就需要某些类型的查找表。

更重要的是，由于这类系统能够显示 1600 万种不同的颜色，甚至 256 级灰度，因此与人眼相比，能识别出更小的渐变。在较好的观看条件下，人眼通常只能分辨几十种不同的灰度级和几百种不同的颜色。从这一方面来看，图像处理系统的显示硬件并不能很好地将图像信息传递到用户。例如，如果图像中的许多像素都是明亮的，那么就无法区分它们。如果图像中存在一些暗像素，那么我们就不可能只是增强图像的线性对比度。相反，存储值和显示值之间需要有一种更为复杂的关系。

对像素亮度的控制，可以用一个把为每个像素存储的亮度值与一个显示值相关联的传递函数来描述。如果这种关系是一对一的，那么对于每个存储值，存在相应且唯一的显示值（尽管视觉上不一定可辨别）。在某些情况下，使用非一对一关系的传递函数有优点，且有必要：几个存储的值用相同的亮度值显示，以便使其他存储值可以进一步展开来增大它们视觉上的差异。

图 5.1(a)显示了一幅图像（干油漆表面，用 SEM 观察），其中 256 个不同像素亮度值不能在计算机显示器上全部被识别；图像的印刷版本必定更糟。第 3 章表明，半色调图像中不同印刷灰度级的数量是由打印机墨点尺寸的变化决定的。本书所用的激光照排机与典型的办公室激光打印机相比，具有更高的分辨率和更多的灰度级，但仍远不如电脑显示器或相片底片。

图 5.1(a)中图像的明亮和暗色区域的细节，甚至在显示器上也看不到（进而在打印件也肯定也看不到）。修改 LUT 可以增大其中一个区域的能见度，或者同时增大暗色区域和明亮区域的能见度，但这样做也要付出代价。图 5.1 显示了对原始图像调整传递函数及存储亮度与显示亮度间的新 LUT 后，产生的几幅新图像。有些新图像是有用的，因为它们揭示了额外的细节，但调整目的主要是为了提升视觉效果。一种非线性关系可以扩张某部分灰度范围，同时压缩另一部分灰度范围。在摄影处理和模拟显示电子学中，这可以通过改变伽马（曝光密度曲线的斜率）值来实现。但计算机可以创建比在暗室中实现的更为复杂的、非线性的任意传递函数。

反转全部对比度范围会产生相片底片效果，有时这会提高细节的可见性。图 5.2 使用一幅 X 射线图像说明了这一效果，X 射线图像通常是通过负片来进行检查的。只反转亮度范围的一部分有时会产生视觉奇效，摄影师称此为负感作用，因此也可用来显示阴影区和饱和区的细节［见图 5.1(h)］。

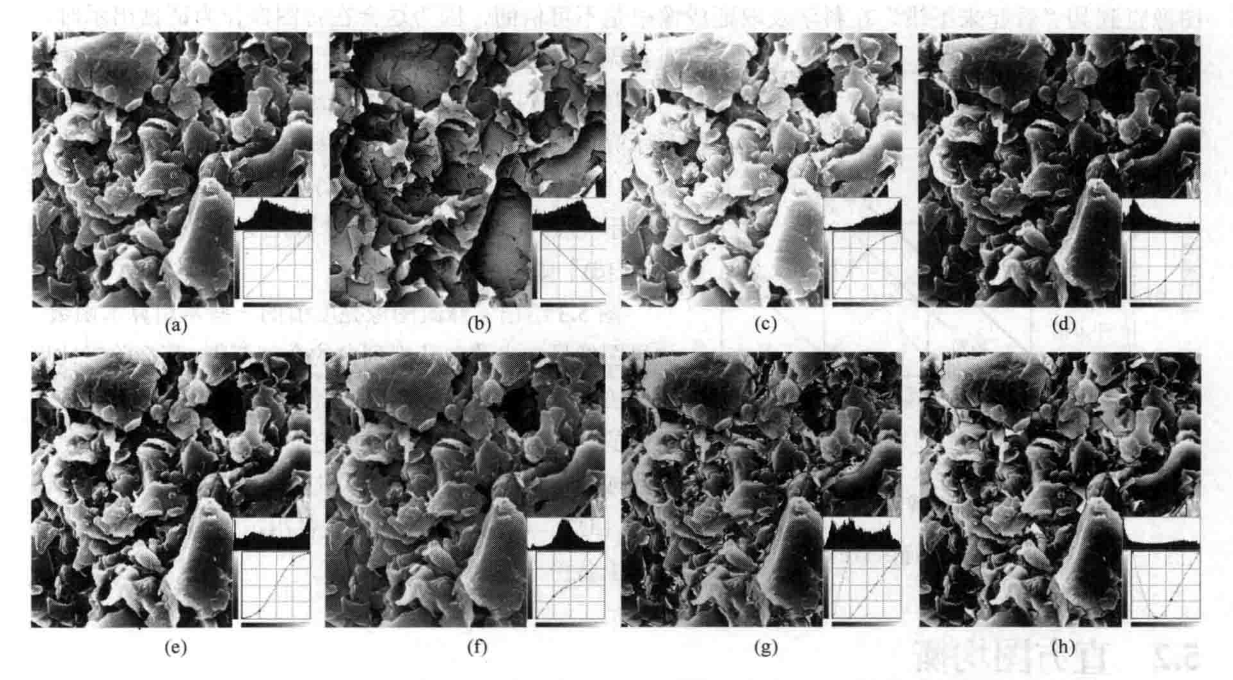


图 5.1　具有全部亮度值范围的原始图像，以及显示传递函数的几个例子，这些传递函数扩展或改变了该范围内不同部分的对比度。每幅图像中右下角的图形中，水平轴表示存储的像素亮度值，垂直轴表示显示的亮度值。同时显示了每幅结果图像的直方图：(a)变换函数是一个恒等函数的原始图像，显示了实际存储的亮度值；(b)反转图像；(c)将伽马亮度增大到中性灰度，在阴影区显示了更多的细节；(d)将伽马亮度减小到中性灰度，在明亮区域显示了更多的细节；(e)对比度增大后的结果；(f)对比度减小后的结果；(g)带状或环绕结果；(h)负感结果

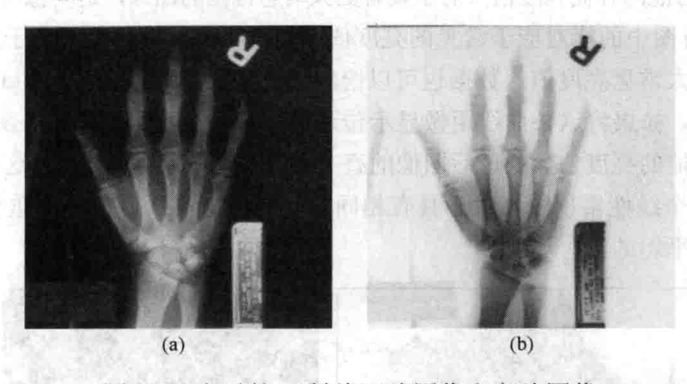


图 5.2　人手的 X 射线正片图像和负片图像

增大传递函数的斜率使其呈环绕状，会生成如图 5.1(g)所示的一幅图像，该图像中几个不同的存储亮度值可能有相同的显示亮度。如果观看者比较熟悉图像的整个组成，那么轮廓线可能不是太具有破坏性，相反它可以增强较小不同的可见性。但随着伪彩色的使用，这类处理很容易过度，进而混淆而不是增强大多数图像。

当然，试验式地修改传递函数，直到图像向观察者最好地显示那些感兴趣特征，提供了一种相当灵活的工具。这通常是人们所选的增强路径，因为这种方法可根据操作的判断，增大某些图像细节并隐藏一些图像细节的能见度。当然，这类似于在暗室中人为控制图像的对比度。多数情形下，我们需要对一系列图像等同地使用可复制和有意义的传输函数，进而进行正确的比较。人为“调整”

图像直到其“看起来不错”在科学及取证成像中是不可信的，因为这会在将图像作为证据出示时，有意无意地存在先入之见。

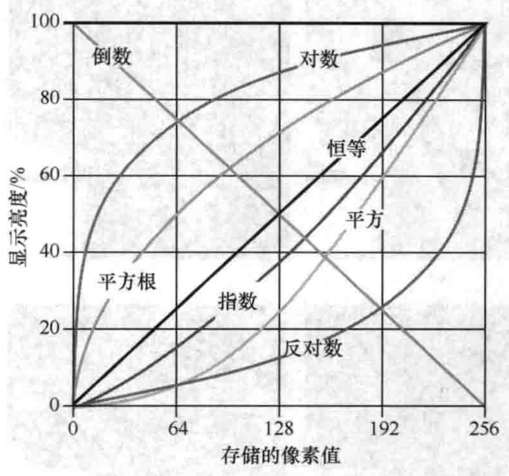


图 5.3　几个显示传递函数的示例

最常见的一类传递函数是遵从某种简单数学关系的显示亮度与存储亮度的关系曲线（如对数曲线或幂律曲线）。上凸伽马曲线［见图 5.1(c)］会压缩刻度范围内亮端的显示亮度，同时扩展暗端的显示亮度。下凹伽马曲线［见图 5.1(d)］则与此相反。

图 5.3 给出了修改图像亮度值的一些常用算术函数。如果图像最初未覆盖从黑到白的全部范围，那么除对比度扩展（将原始范围扩展到显示器的整个范围）外，这些函数中的任意一个还可用在其他方面。这些传递函数的值的曲线或表可预先计算并存储，以便可快速载入，进而修改显示 LUT。可在手边放几个表备用，如为伪彩色显示在磁盘上备份一系列彩色 LUT。

5.2　直方图均衡

除标准数学函数外，有时需要为特定图像构建传递函数。与上面给出的任何函数不同，构建传递函数的目的是提供可重复并有最优结果的一种专用算法。最常用的这类方法是直方图均衡（Stark & Fitzgerald，1996）。为理解什么是图像均衡，我们可从了解图像亮度直方图开始。

图 5.4 显示了一个例子，其中所用的图像与图 5.1(a)相同。传统直方图会显示图像中的许多像素，这些像素都有 256 个可能的存储亮度值（对于具有更大动态范围的图像，这些像素的亮度值位于 256 个线性计算条中）。直方图中的峰对应于常见的亮度值，而常见亮度值可能对应于所出现的特殊结构。谷对应于图像中的不太常见亮度值。数据也可以绘制成一条累积曲线［图 5.4(a)中显示为红色］，即这些值的积分或求和。如果将这条曲线用做显示传递函数，那么显示结果［见图 5.4(b)］中，会机会均等地使用 256 个可能的亮度值。处理后图像的直方图［见图 5.4(c)］显示了这一均匀分布（因此称为直方图均衡）和一个线性累积图（由于具有相同初始亮度值的所有像素被重新赋值为同一新值，因此直方图只是近似平坦的）。

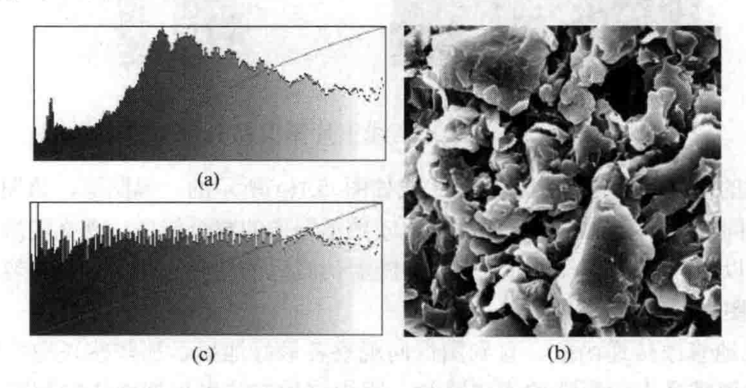


图 5.4　直方图均衡：(a)图 5.1(a)中原始图像的直方图，累计总和显示为红色；(b)应用均衡函数后的图像；(c)处理后的直方图，显示了均匀分布的值和累积曲线

通常，每幅图像都会有其唯一的亮度直方图。甚至同一标本或场景图像的不同区域，也具有不同的直方图，具体取决于每幅图像所覆盖的区域。改变整体光照或相机设置会移动直方图中的峰。此外，许多真实图像会展示出某些特征内的亮度变化（如从边缘到中心）。从有效地利用显示器上可用亮度级的观点来看，有些值未被充分利用。直方图均衡的基本原理是展开峰区中显示的亮度级，并在谷区压缩它们，进而使显示器中相同数量的像素显示每种可能的亮度级。

基于图像直方图重新分配像素的亮度值，直方图均衡实现了这一重新分布。各个像素保留了其亮度顺序（即它们仍然亮于或暗于其他像素，但具有几个初始值的像素例外，这些像素被分配了相同的均衡值）但值被改变，因此平均而言，等同数量的像素都有一种可能的亮度值。多数情形下，这会展开不同区域相交区域中的值，用高亮度梯度显示区域中的细节。

像素具有类似亮度值的几个区域构成的图像，其直方图中会有峰出现。这些峰的大小给出了不同相区的相对面积，这对图像分析非常有用。对图像进行直方图均衡会展开这些峰，同时通过将相同或非常接近的亮度值赋给那些具有中等亮度的少量像素，来压缩直方图的其他部分。这种均衡可让我们看到原始图像中看起来相同区域内的微小变化。

过程很简单：如果亮度值的范围是 0～255，那么对于原始图像（及其直方图）中的每个亮度级 j，新赋的值 k 计算为

$$k = 256 \cdot \sum_{i=0}^{j} \frac{N_i}{T} \tag{5.1}$$

式中，求和（通过对直方图积分）用于计算图像中亮度等于小于 j 的像素的数量，T 是像素总数（或直方图下的总面积）。

图 5.5 给出了直方图具有明显峰和谷的一幅图像的例子。原始金相试样有三个相区，这三个相区分别有黑色、中等和亮灰色值。直方图均衡会展开峰中的值，即增大像素间的差别以便看起来更清楚。因此，原始图像中不可见明亮区域内的阴影现在变得可见了。因为原始图像未覆盖从黑到白的完整范围，且仅包含 8 位数据，因此结果直方图［见图 5.5(d)］中出现了许多间隙，尽管累积曲线图表明值的一般分布已基本均匀。此外，一些原本具有不同值的像素被分配了相同的值。

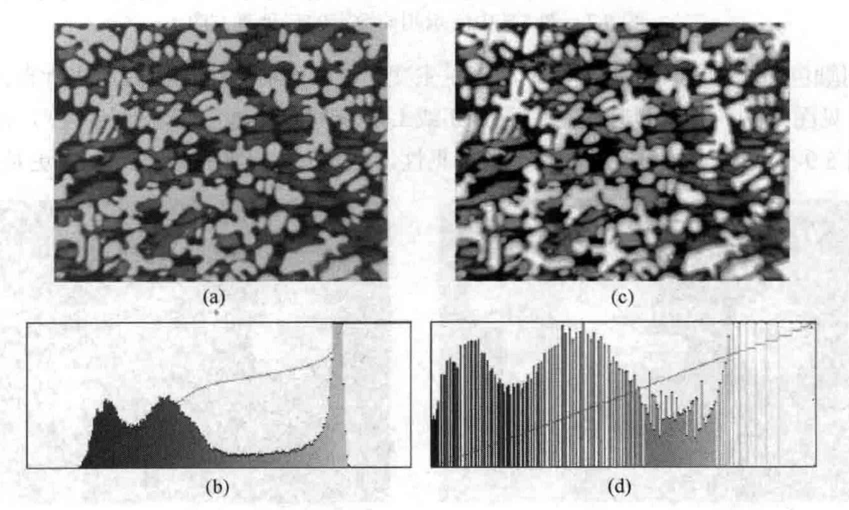


图 5.5　应用到一幅 8 位灰度图像的直方图均衡：(a)原始图像；(b)显示了峰和谷且未完全利用亮度范围的直方图；(c)均衡后的直方图；(d)结果直方图

对于彩色图像，可显示与三个颜色轴或通道对应的三个直方图。如图 5.6 所示，这可以对 RGB、

HSI 或 Lab 彩色坐标进行。但这些直方图中的任何一个都不能完全表征图像，因为它们不能显示相同像素的组合值。图 5.7 给出了一个三维直方图，其中直方图中的点具有对应于彩色值的坐标，彩色值则显示了值的每种组合的像素数量。较暗的值表示具有通道值的特定组合的大量像素。还给出了三维直方图在立方体的每个二维面上的投影。第 7 章在介绍为阈值处理选取颜色组合时，会用到这种方法；第 10 章中在介绍测量不同颜色通道（如不同荧光染料）中亮度的共位时，也会用到该方法。

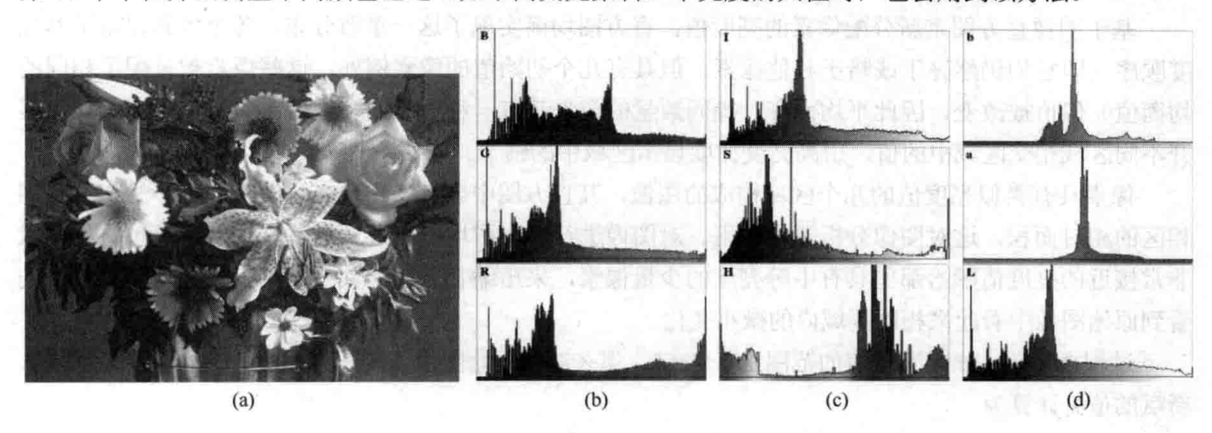


图 5.6　一幅彩色图像中各通道的直方图：(a)图像；(b)红、绿和蓝直方图；(c)色相、饱和度和亮度直方图；(d)L、a 和 b 直方图。注意图(c)中的亮度通道和图(d)中的 L 通道不同，因为它们有着不同的权重

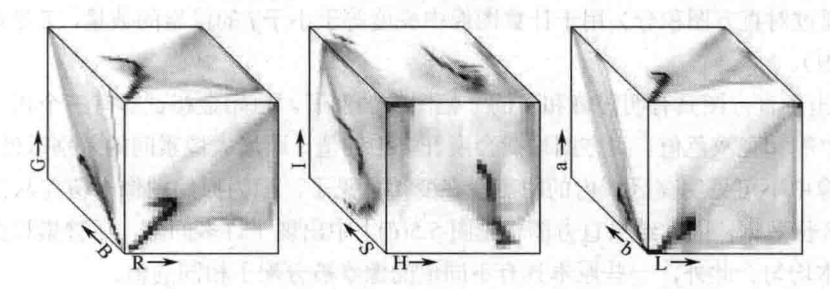
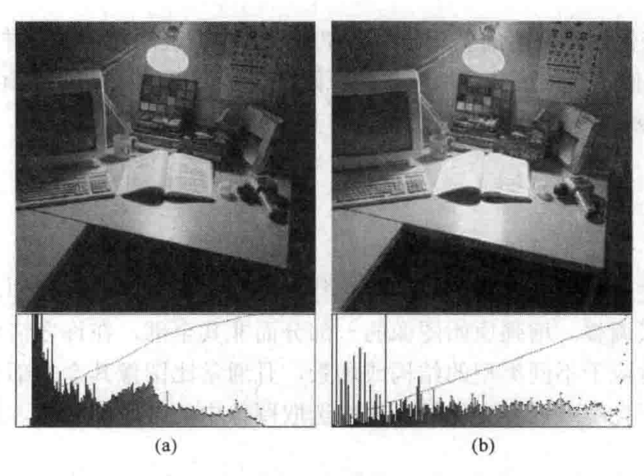


图 5.7　图 5.6 中所示相同数据的三种直方图

由于不同颜色的比例在图像的不同区域会发生变化，因此对各个颜色通道执行直方图均衡会产生离奇的结果（见图 5.8）。正确步骤是工作于 HSI 或 Lab 空间，让颜色信息保持不变，而只处理亮度或灰度值。如图 5.9 所示，这会明显增强细节的可见性，对具有大范围亮度值的图像更是如此。



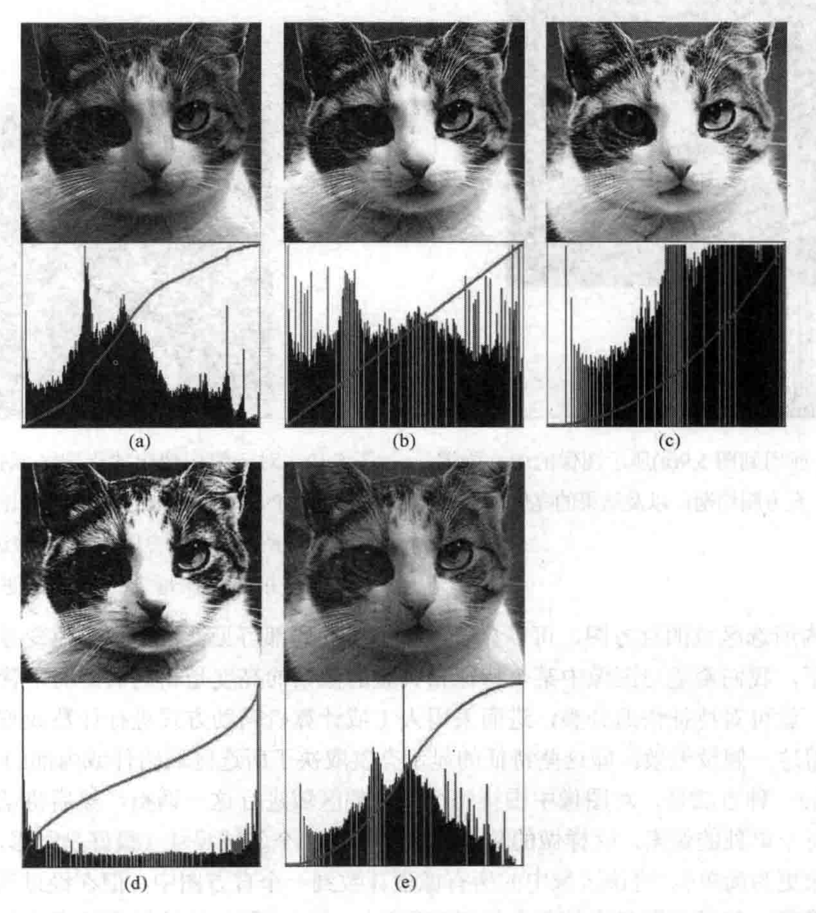
图 5.8　(a)对图 5.6(a)中所示图像的 R、G、B 通道应用直方图均衡的结果；(b)仅对亮度应用直方图均衡而保持彩色信息不变的结果



(a) (b)

图 5.9 对彩色图像的亮度通道应用直方图均衡：(a)原始图像；(b)处理后的结果

均衡只是直方图形状调整的一个例子。人们也会用到其他的预定形状，如根据人类亮度感知来均衡值的形状（Frei，1977）。图 5.10 比较了几个不同的函数，由累积直方图的形状，我们很容易就可理解这些函数。图 5.11 显示了对图 5.9 所示图像的亮度通道应用中心强调均衡后的结果。重新分配一幅图像中的亮度值来调整直方图的形状，进而与另一幅图像的直方图匹配是可能的，但这仅在少数情形下有用，譬如匹配在不同时间或不同照明条件下拍摄的同一场景的两张照片。



(a) (b) (c)

(d) (e)

图 5.10 直方图修改（显示了每幅图像及其直方图，以及以红色显示的累积直方图）：(a)原始图像；(b)线性均衡；(c)指数函数（亮度强调）；(d)幂律函数（两端强调）；(e)误差曲线函数（中心强调）

多数情形下，一些像素亮度值的重新分配会使直方图中出现间隙，并合并一些值使得最初亮度不同的像素被赋予相同的值。使用更高位深图像的原因之一是，可以在 256 值直方图中不出现间隙，或结果图像中不出现轮廓线或色调分离的情形下，对其亮度值进行调整。

5.3　局部均衡

任何直方图操作过程都不需要对整幅图像进行，不管是简单的线性拉伸、预定数学传输函数的应用，还是直方图形状调整。增强原始图像的一部分而非其全部，在许多情形下也是有用的。当某幅图像的几个大区域对应于不同类型的结构或场景，且通常比图像其余的部分更亮或更暗时，尤其如此。在可人为或基于亮度或内容变化的算法来选取图像中的某些部分时，就可使用一种选择的均衡来突出局部细节。

图 5.12 显示了图 5.1(a)所示图像的区域，其中分别选取和修改了几个区域。两个手工画出的区域分别是任意矩形和椭圆形。其他区域则是结构（颗粒或孔洞）的轮廓，能从图像其他部分分离出它们的原因是，从局部来看它们的边界得到了清晰的定义。定位边界边缘的方法将在本章后面讨论。每个区域基于该区域内像素的直方图单独处理，产生更大的局部对比度并增加了细节能见度。

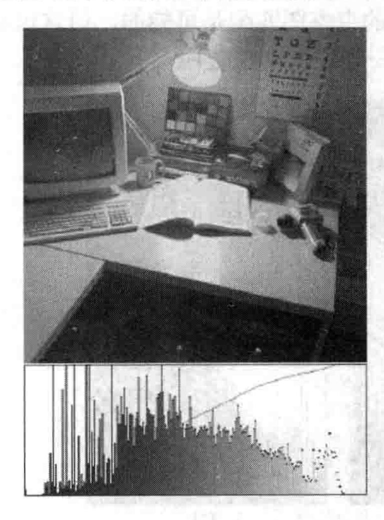


图 5.11　应用到图 5.9(a)所示图像的中心强调直方图均衡，以及结果的亮度直方图

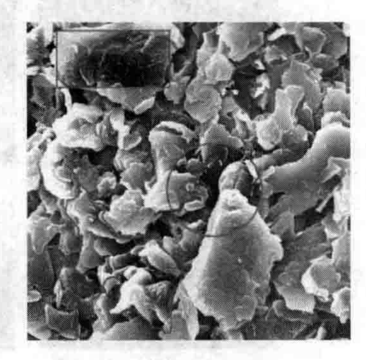
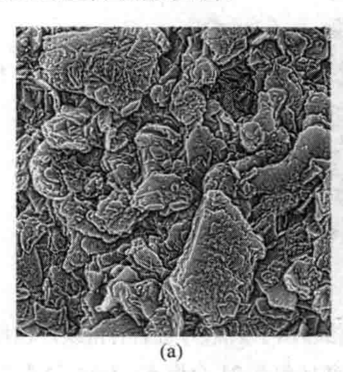


图 5.12　对一幅图像中的指定区域执行直方图修改的选择性对比度增强。在该例中，图 5.1(a)所示图像中的几个区域已被选取（红色圈出），并分别在每个区域内部进行直方图均衡

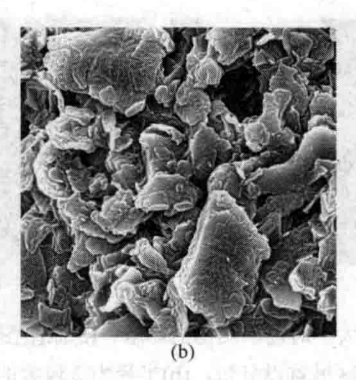
修改图像内所选区域的直方图，可以显著改善细节的局部可见性，但也会改变亮度和结构间的关系。多数情形下，我们希望与图像中某个特征相关联的像素的亮度是相同的，而不管该像素在图像中的何处。这样，就可对特征快速分类，进而采用人工或计算机辅助方式进行计数或测量。灰度关系的局部修改会使得这一假设无效，即这些特征的显示亮度取决于所选区域附件或内部的其他特征。

更为可取的一种方法是，对图像中围绕每个像素的区域进行这一调整，然后将结果单独应用基于其周围邻域的每个单独的像素。这样做的基本方法是指定一个邻域尺寸（最好是圆形，但通常是方形，因为其实现起来更为简单）。将该区域中的所有像素计数到一个直方图中，但在经过调整步骤后，新值仅应用到中心像素。然后对图像中的每个像素应用这一处理，即总是使用原始亮度值来构建每个区域的直方图。

图 5.13 显示了一个例子，其中已对图像中的每个像素应用了局部直方图均衡。半径为 6 像素的一个圆形邻域以图像中的每个像素为中心，而邻域内的 137 个像素的直方图则用于执行均衡操作。但新像素值只为中心像素保留，然后使用它来形成一幅新图像。实际计算很简单，因为对于每个像素来说，均衡后的亮度值正好是在邻域中较暗像素的数量。由于邻域包含 137 像素，该处理会生成一幅新图像，图像中在整个亮度范围内展开的不同亮度值的最大数是 137 而非原来的 256，但这一数字仍然超过了显示器中可视觉分辨的数量。即使对于一个小邻域（如包含 21 个像素的 5 像素宽的圆），灰度值的数量通常也足以应付视觉检查。



(a)



(b)

图 5.13　应用到图 5.1(a)所示图像的局部均衡。局部细节被增强，同时抑制了大规模的对比度：(a)局部均衡结果；(b)1/3 局部均衡结果与 2/3 原始图像的组合结果

在例子中明显可以看到，局部均衡处理仍会使得稍高于周围像素的像素稍亮（甚至觉察不到的），反之亦然。这增强了点和边缘附近的对比度，即揭示了亮与暗区域中的细节。在近乎一致的区域，这会产生虚假的对比度变化，即放大图像中噪声伪影的能见度。总之，在进行增强运算之前，使用第 4 章中介绍的方法消除或降低噪声是重要的。

例子图像还表明，局部均衡的结果降低了图像中的整体或大规模对比度。明亮颗粒表面和暗色孔洞均被约化到一个平均灰度，进而隐藏了结构的许多整体形状。我们有时并不期望这种约化，将某个原始图像的某个百分比加到处理结果中，可补偿这种约化，如图 5.13(b)所示。在其他情形下，如图 5.14 所示，我们期望能消除大规模的亮度变化。当然，也可采用混合局部均衡结果与全局直方图均衡结果的方法来消除大规模的亮度变化（Woon & Song，2007）

为便于比较，图 5.14 还显示了经背景调匀来消除图像中大规模对比度的结果。使用一个灰度开运算（先腐蚀，即每个像素用其最亮的邻点代替，然后膨胀，即每个像素用其最暗的邻点代替，如第 4 章所述）消除了指纹标记。两个排序运算都使用了一个半径为 7 像素的圆形邻域。该背景然后除以原始图像来生成结果。一般来说，与局部均衡法相比，这种方法会产生视觉上更粗糙的结果。

改变局部处理所用邻域的大小可对该处理提供一些控制，如图中所示。只要邻域是大到足以涵盖所示纹理规模，且小到足以抑制亮度的长程变化，所得结果通常就会令人满意（Wu et al.，2005）。其他更有效（且通常结合术语“自适应”使用的）改进是，按照像素与邻域中心的远近，或按照像素亮度与中心像素亮度的类似程度，或包含邻域中只与中心像素类似的像素，对这些像素分配权重。

增强细节可视性的另一种局部均衡技术是方差均衡（Laine & Huda，2000)。这种技术也使用了一个移动邻域，邻域中仅进行修改中心像素的计算。计算区域中像素的统计方差，并与整幅图像的方差进行比较，然后调整像素值的大小来使得局部方差与全局方差匹配。结果同样会增大均匀区域的对比度，如图 5.15 所示。硬币表面上的压痕和划痕在处理后更加明显。将增强的对比度应用到渲染表面重建出表面几何形状，我们能更清楚地看到表面标记的表面增强，如第 15 章中的图 15.18 所示。

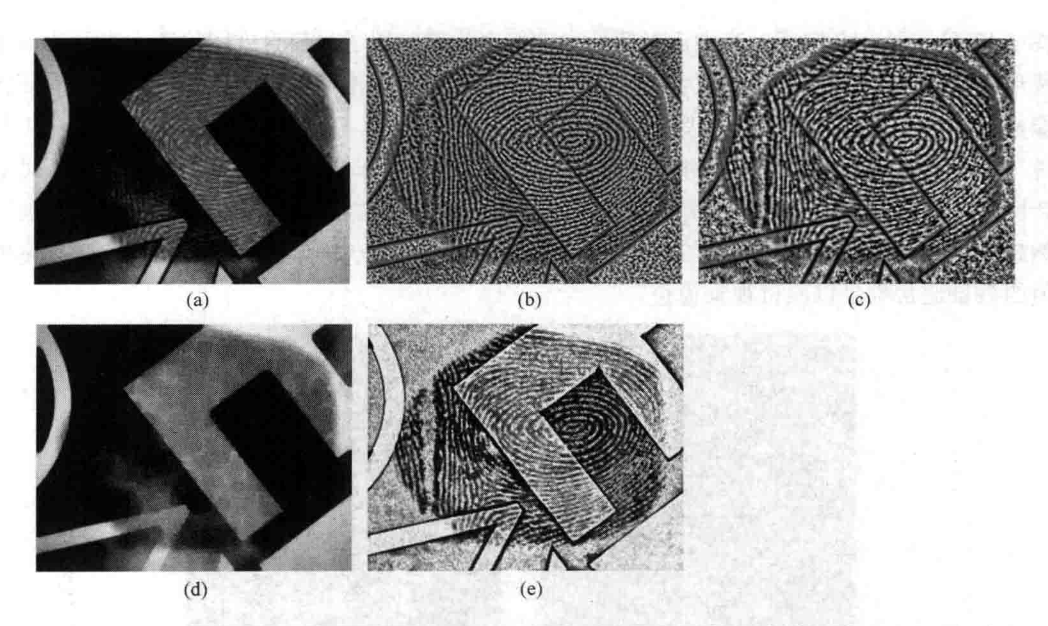


图 5.14　邻域大小对局部均衡的影响：(a)原始图像显示了杂志封面上的指纹，该指纹在明与暗的印刷区域难以分辨；(b)半径为 2 像素的邻域包含 21 个像素；(c)半径为 7 像素的邻域包含 177 个像素。两个结果均抑制了来自背景的对比度，并使得指纹中的脊纹标记更为明显。为便于比较，显示了使用灰度腐蚀和膨胀获得的结果；(d)使用开运算删除了脊标记；(e)图(a)与图(d)相除的结果

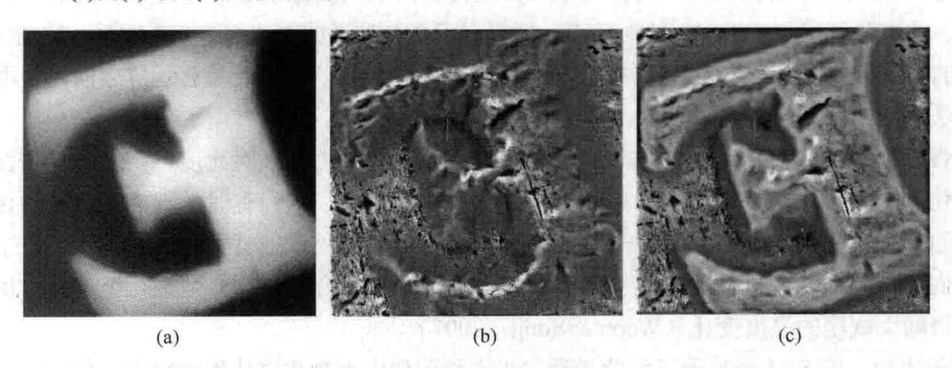


图 5.15　应用到表面图像的局部增强：(a)原始图像（上有凸起字母的硬币）；(b)应用局部均衡后的结果；(c)应用方差均衡后的结果

处理彩色图像时，正确的方法是仅对亮度数据应用局部均衡。首先，将原始的 RGB 格式图像转换到 HIS 或 Lab 空间，然后修改亮度值，再后与原始色相和饱和度值组合，以便计算新的 RGB 值，进而能够显示结果（Buzuloiu et al.，2001）。

5.4　拉普拉斯算子

图像对比度的局部均衡会增强边界处的局部对比度，如图 5.16(b)所示。局部对比度增强的结果是，边界看上去更清晰（尽管例子中近乎均匀区域内的增大的噪声抵消了这一增强）。边缘增强还有其他几种方法，与上面讨论的均衡相比，这些方法对整体亮度级、噪声及所示细节类型或大小的影响较小。

第一组运算使用具有乘法核的邻域（类似于第 4 章中用于噪声平滑的邻域）。在关于平滑的那一节中，卷积核写为一个整数数组。例如，

$$\begin{bmatrix} +1 & +2 & +1 \\ +2 & \mathbf{+4} & +2 \\ +1 & +2 & +1 \end{bmatrix}$$

这个 3×3 核的含义是，中心像素亮度值乘以 4，中心像素上下左右的四个像素的亮度值乘以 2，而四个从对角像素的亮度值乘以 1。总值相加，然后除以 16（9 个权值之和）来为中心像素产生一个新像素亮度值。也可使用其他权重数组，有些数组要远大于上面的 3×3 数组。对于平滑，所有核均关于中心对称（至少它们在方形像素网格和整数值的约束下近似对称），并且多数核只有正权重值。

大致对称且出现负值的一个非常简单的核，是经典的 3×3 拉普拉斯算子：

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & \mathbf{+8} & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

这一操作的实质是，从 8 倍的中心像素亮度值中减去 8 个邻近像素的亮度值。因此，在图像中亮度均匀的区域或具有均衡亮度梯度的区域，应用该核的结果就是将值降到零。当形为点、线或边缘的邻域内出现变化时，拉普拉斯算子的结果是一个非零值。它可能为正，也可能为负，具体取决于中心点相对于边缘的位置等。

出现正值和负值时，为显示结果，通常会添加一个中间灰度值（当图像的灰度值范围是 0～255 时，所添加的值为 128），使得零点为中间灰度，进而便于观察由拉普拉斯算子产生的亮值和暗值，如图 5.16(c)所示。有些程序则会画出结果的绝对值，但会在沿边缘处产生双线，而这会误导观看者及后续的处理与测量操作。

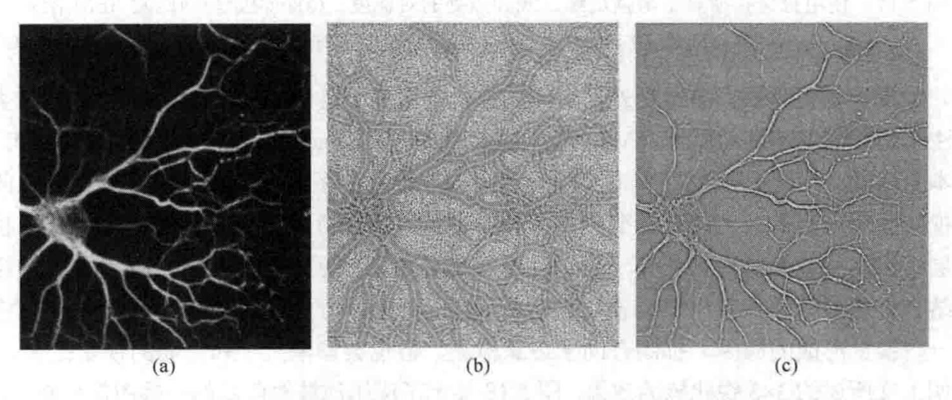


图 5.16　细节的增强：(a)原始图像（树突细胞）；(b)局部均衡结果显示了细小结构，但增大了均匀区域内的噪声；(c)拉普拉斯结果增大了沿边缘的对比度，而噪声放大很少

顾名思义，拉普拉斯算子是对亮度 B 在 x 和 y 方向的二阶微分算子的逼近：

$$\nabla^2 B \equiv \frac{\partial^2 B}{\partial x^2} + \frac{\partial^2 B}{\partial y^2} \tag{5.2}$$

它具有旋转不变性，因此对不连续出现的方向不敏感。它会突出图像中的点、线和边缘，抑制均匀和平滑变化的区域，结果如图 5.17 所示。这幅拉普拉斯图像本身失去了目标整体形状的信息。将边缘的拉普拉斯增强加到原始图像，可恢复观看者能解释的整体灰度变化。它还会通过在不连续处局部增大对比度来锐化图像，如图所示。这种效果的实现方法是，改变核中的权重（并消除 128 的补偿）使其变为

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & \mathbf{+9} & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

该核通常称为锐化算子，因为它会在边缘处增大图像的对比度。该处理过程有两种不同的解释。首先，使用扩散处理对一幅模糊图像建模，扩散处理会将亮度展开到整幅图像，它服从偏微分方程

$$\frac{\partial f}{\partial t} = k \cdot \nabla^2 f \tag{5.3}$$

式中，模糊函数是 $f(x, y, t)$，而 t 表示时间。如果将其按时间进行泰勒级数展开，则不模糊图像可表示为

$$B(x, y) = f(x, t, \tau) - \tau \frac{\partial f}{\partial t} + \frac{\tau^2}{2} \frac{\partial^2 f}{\partial t^2} - \cdots \tag{5.4}$$

忽略高阶项，上式变为

$$B = f - k \cdot \tau \nabla^2 f \tag{5.5}$$

换言之，图像 B 可通过组合拉普拉斯运算（乘以一个常数）与模糊图像来恢复。虽然模糊图像建模为扩散处理是最佳逼近，且比例常数是未知的或任意的，但这至少为这一方法给出了某些可信度。

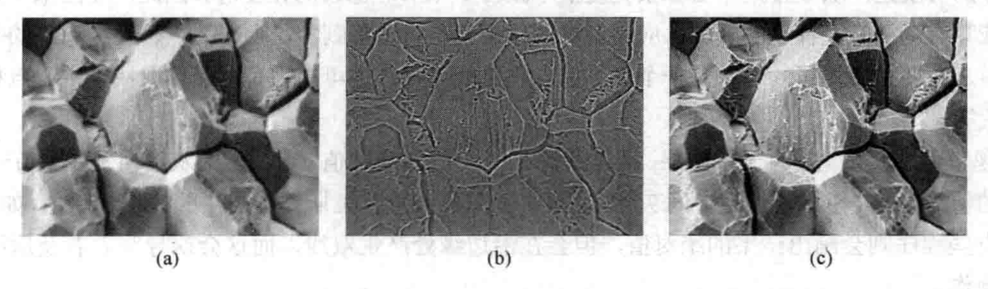


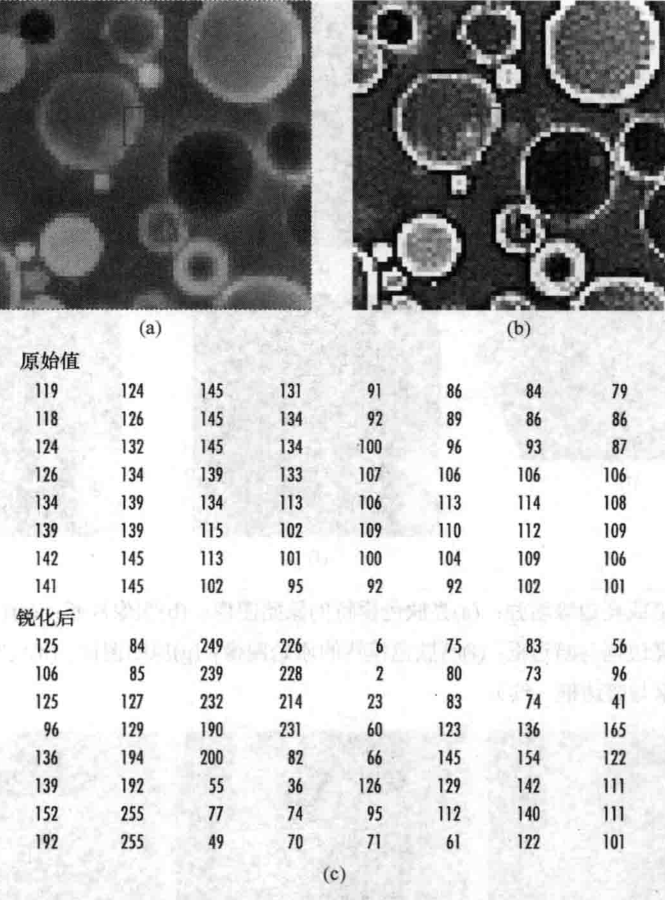
图 5.17　使用拉普拉斯算子增强边缘、线和点处的对比度：(a)陶瓷断裂的原始 SEM 图像；(b)应用拉普拉斯算子后的结果；(c)拉普拉斯算子运算结果与原始图像相加后的效果

同样，如果不是更重要，事实是处理后的图像“看起来不错”。如第 2 章所指出的，人类视觉系统本身会关注于边缘而忽略均匀区域（Marr & Hildreth，1980；Marr，1982；Hildreth，1983）。抑制是发生在眼睛本身内的一个过程。视网膜内部神经元间的连接，会按照其邻域来抑制一个区域的输出。这种能力是视网膜所固有的。直接连接到视网膜传感器的是两层处理神经元，其功能类似于拉普拉斯运算。第二层中的水平细胞会对来自第一层中几个相邻传感器的信号进行平均，而无长突神经细胞与双极细胞会结合该信号与原始传感器的输出。这种局部抑制有助于提取边界和边缘。此外，它也说明了第 2 章中一些错觉形成的原因。抑制有助于边缘检测，但也会影响大小和方向的视觉比较。

为说明上文所述的 3×3 锐化核的效果，图 5.18 显示了应用滤波器前后的一幅图像片断（已放大到可显示单个像素）和实际像素值（在 0～255 范围内显示为整数）。滤波器的作用是使得边缘暗侧的像素仍较暗，而亮侧的像素仍较亮，以此增大边缘对比度。

取决于局部区域是暗还是亮，该技术的一种改进可用来将边缘增强仅局限到较亮或较暗的一侧，这可改善结果的视觉外观，如图 5.19 所示。其中比较了亮与暗的肤色模型。多数情形下，明亮细节在深色背景上看起来最好，反之亦然。这种相同的选择性增加可与下面介绍的反锐化掩模及高斯差滤波器一起使用。

在该例中，还要注意滤波器只应用于图像的亮度通道，而颜色信息不变。这对于所有图像处理运算都很重要，因为对原始红色、绿色和蓝色通道的运算（Wu et al.，2008）会产生与值成比例的变化，而这些变化会改变颜色，并严重退化处理后的图像外观。图 5.20 说明了这一差异。



(a) (b)

原始值

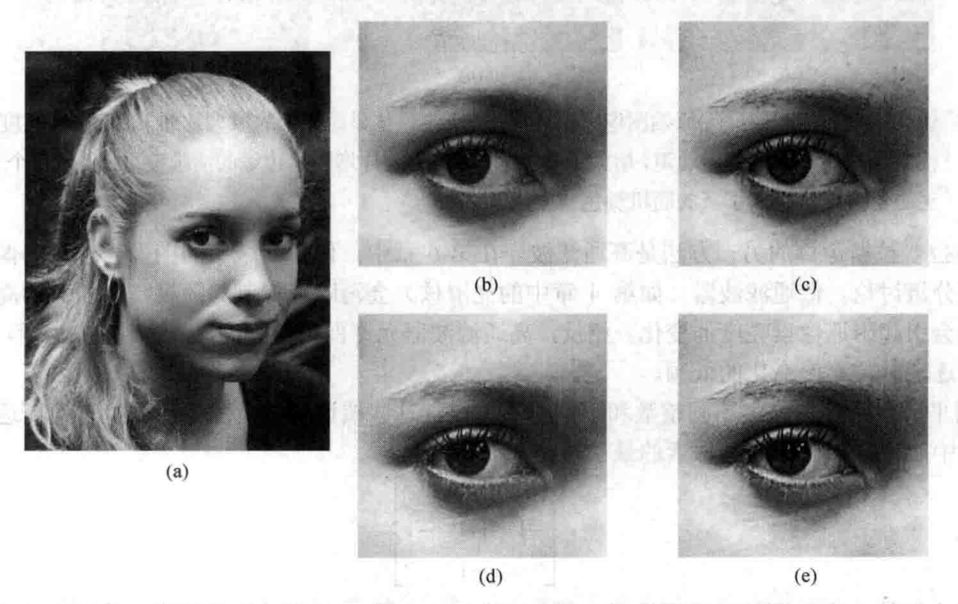
119	124	145	131	91	86	84	79
118	126	145	134	92	89	86	86
124	132	145	134	100	96	93	87
126	134	139	133	107	106	106	106
134	139	134	113	106	113	114	108
139	139	115	102	109	110	112	109
142	145	113	101	100	104	109	106
141	145	102	95	92	92	102	101

锐化后

125	84	249	226	6	75	83	56
106	85	239	228	2	80	73	96
125	127	232	214	23	83	74	41
96	129	190	231	60	123	136	165
136	194	200	82	66	145	154	122
139	192	55	36	126	129	142	111
152	255	77	74	95	112	140	111
192	255	49	70	71	61	122	101

(c)

图 5.18　3×3 锐化滤波器运算：(a)原始图像片断（含气泡的环氧树脂）；(b)应用滤波器后，边缘由亮和暗边框描绘；(c)红色圈定区域的像素值。边缘暗端的像素值减小，而亮端的像素值增大



(a) (b) (c) (d) (e)

图 5.19　只允许暗或亮边缘增强：(a)亮肤色模特的原始图像；(b)图像片断；(c)仅暗边框；(d)仅亮边框；(e)亮边框与暗边框；(f)暗肤色模特的原始图像；(g)原始图像；(h)仅暗边框；(i)仅亮边框；(j)亮边框与暗边框

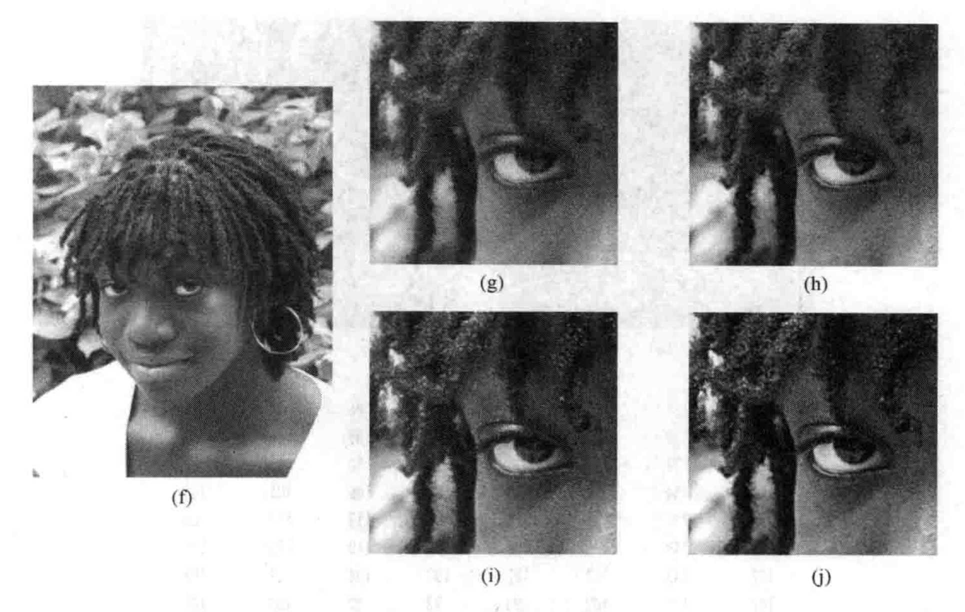


图 5.19　只允许暗或亮边缘增强：(a)亮肤色模特的原始图像；(b)图像片断；(c)仅暗边框；(d)仅亮边框；(e)亮边框与暗边框；(f)暗肤色模特的原始图像；(g)原始图像；(h)仅暗边框；(i)仅亮边框；(j)亮边框与暗边框（续）

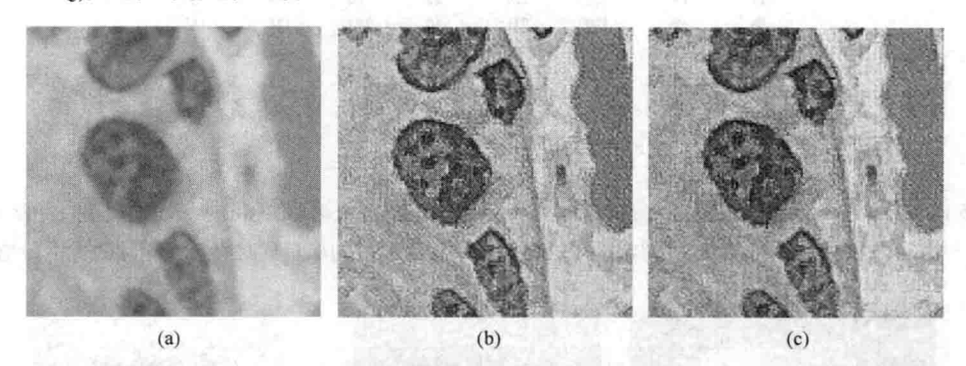


图 5.20　锐化一幅彩色图像：(a)原始图像［图 4.5(a)所示染色体显微图像的放大片断］；(b)处理亮度通道而不处理色调或饱和度通道，增大了噪声的可见性，同样增加了边缘的清晰度；(c)处理各个 RGB 通道给各像素引入了新的随机颜色

描述拉普拉斯运算的另一方法是高通滤波。在第 6 章中，傅里叶域中的图像处理按图像亮度的高频和低频分量讨论。低通滤波器（如第 4 章中的光滑核）会消除或抑制与随机噪声相关的高频变化，因此可能会引起附近像素亮度的变化。相反，高通滤波器允许保留这些高频（即通过滤波器），同时消除对应于逐渐整体亮度变化的低频。

如同平滑核，存在许多不同整数和不同大小的核，这些核可用来对图像应用拉普拉斯运算。仅使用 4 个与中心像素同侧的相邻像素的最小核是

$$\begin{bmatrix} & -1 & \\ -1 & \mathbf{+4} & -1 \\ & -1 & \end{bmatrix}$$

使用小核的一个问题是差异严格限于局部，即把中心像素与其近邻相比较。最常见的噪声类型由邻近像素间的随机变化组成，即使是在图像中看起来均匀的区域，因此这种这种方法可使它更加明显。

提高细节和边缘可见性的一种更常用方法是反锐化掩模。对于使得图像看起来更清晰的技术来说，

"反锐化掩模"一词看起来很奇怪。它源自传统摄影暗室中所使用的一种方法。第 1 章中的图 1.30 说明了其步骤。首先，将原始底片放到胶片上，以 1:1 的放大率和轻微的失焦或模糊，完成接触式印片。胶片显影后，就生成了一个新底片，然后将两个底片对齐并夹在一起。原底片上的明亮区域被胶片上的暗色区域覆盖（反锐化掩模），几乎不允许光穿过。仅打印出与原始底片不匹配的轻微失焦区域。这有些类似于拉普拉斯运算，拉普拉斯运算从原图中减去平滑后（失焦）的图像来抑制渐变而通过高频或边缘，但它允许模糊平滑地扩展到更大（且可调节）的距离。

反锐化掩模的计算机实现通常对图像的副本应用高斯模糊，然后再从原图中减去它。如同对拉普拉斯算子一样，结果通常加回原图，以强调边缘和细节，同时保留一些整体的图像对比度。图 5.21 显示该处理增大围绕线和边缘的局部对比度的方式。应用反锐化掩模算子可增加细节的可见性，同时抑制亮度的整体变化。图 5.22 显示了一个例子，它使用了图 5.2 所示的 X 射线图像。在原始图像中，手指中的骨骼较少，因此不像手腕上的骨骼那样密，而且书写标签很不明显。而在处理后的图像中，这些内容更容易看到。

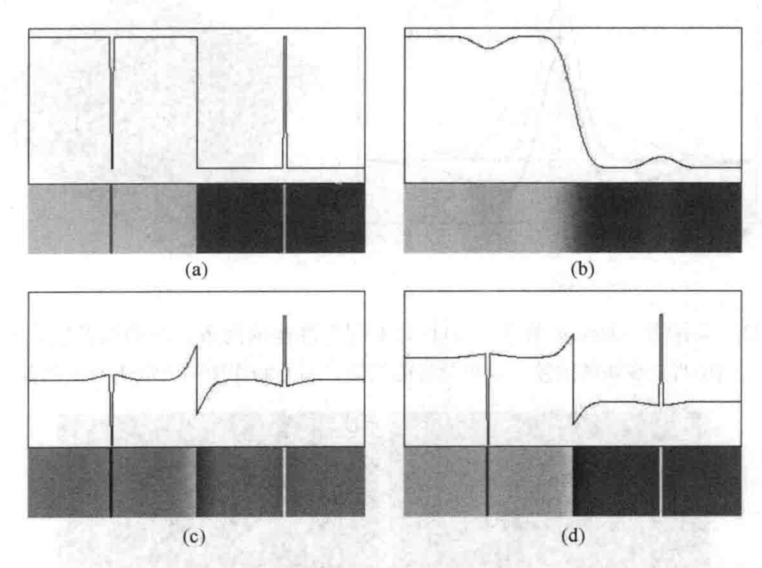


图 5.21　反锐化掩模运算及亮度剖面：(a)原始图像；(b)高斯平滑结果；(c)从原始图像(a)中减去平滑结果(b)后的图像；(d)将差(c)加到原始图像(a)后的结果

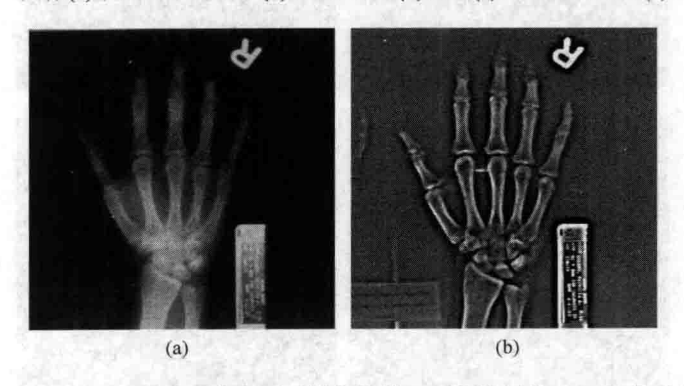


图 5.22　对人手 X 射线图像应用反锐化掩模：(a)原始图像；(b)处理后的图像

与反锐化掩模密切相关，这一技术的通用形式是，从一幅具有不同平滑度的图像中减去该图像平滑后的版本。这称为高斯差（DoG）方法，有人（Marr，1982）认为该方法类似于人类视觉系统对边界和其他特征的定位方式。DoG 实际上在形状上也与用到的另一个函数一致，即高斯拉普拉斯算子或 LoG（也称为 Marr-Hildreth 算子）。

DoG 技术在其抑制图像中随机像素噪声变化的能力上，优于反锐化掩模。使用具有适当但较小标准差的一个高斯核平滑图像，抑制高频噪声，进行第二次高斯平滑（具有更大的标准差，一般是第一次时的 3～6 倍）消除噪声和重要的边缘、线和细节。然后仅保留两幅图像间那些大小们于两个算子间的结构（线、点等）的差。如果说拉普拉斯算子和反锐化掩模可视为高通滤波器，那么 DoG 则可更恰当地视为带通滤波器，因此它会保留所选范围的频率。

具有不同标准差的两条高斯曲线及它们的差如图 5.23 所示；当标准差较小时，该差类似于拉普拉斯算子的剖面，而在标准差小到可忽略时则类似于反锐化掩模的一个剖面。因为这些核的形状以等角视图绘制，因此有时也称之为墨西哥帽或草帽滤波器；这一名称通常是为中心有多个正权值像素的核保留的。图 5.24 比较了一个 3×3 的基于拉普拉斯算子的锐化滤波器、一个反锐化掩模和一个 DoG 滤波器对一幅带噪图像处理的结果。

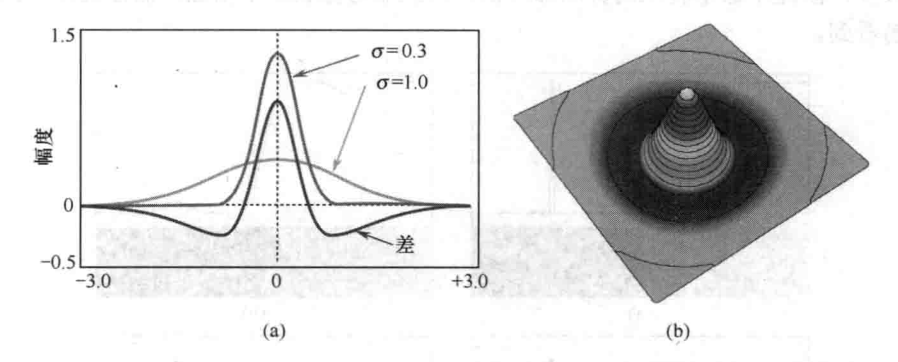


图 5.23 高斯差（DoG）算子：(a)具有不同标准差的两条高斯曲线及它们的差；(b)两个高斯核的差，以等角视图绘制。中心峰周围的“壕沟”值包含负值

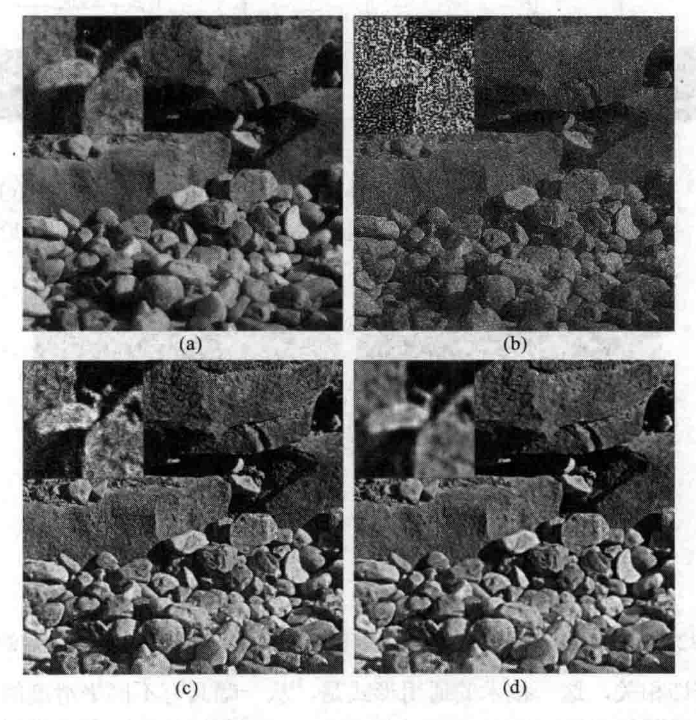


图 5.24 应用到一幅测试图像上的高斯差：(a)原始图像，即第 4 章中的图 4.9(a)；(b)3×3 锐化滤波器；(c)反锐化掩模，其中 $\sigma = 6.0$；(d)使用 $\sigma_1 = 6.0$ 和 $\sigma_2 = 1.0$ 像素的高斯差

以往用于卷积的核，一直相对较小，它由整数权值组成，以使算术运算尽可能快。构造负值之和与正值之和相等的这样一个核可避免整体亮度偏移（进而使特征和边缘方向等同增强，而与它们的方向性无关），并且符合高斯曲线的形状。本书的前几版中包含了几个例子。幸运的是，随着支持浮点运算能力的快速计算机的出现，这并不算是问题。需要时可在观察增强后图像的同时，交互地构建合适的滤波器。选择性地应用上面给出的相同选项，可应用于一个阶的亮侧或暗侧。对灰度或亮度通道的滤波操作仅限于色样－饱和度－亮度空间或 Lab 空间，而不能用于 RGB 空间。

5.5 导数

拉普拉斯算子、反锐化掩模和 DoG 滤波器适用适用于通用目的的视觉增强，但并非是所有情形下的最佳工具。当图像细节有已知的一个方向时，就属于这样一种情形。这样的例子包括色谱制备（蛋白质在电场中沿着泳道延展（见图 5.25）］或者岩芯的年轮图案（见图 5.26）。在重要变化的方向，对这样的一幅图像应用一阶导数，可增强小台阶和其他细节的可见性，如图所示。当然，对于具有有限像素的数字化图像，则不能进行连续求导。相反，相邻像素间的差值计算为有限导数。这种差值在某种程度上是噪声，但在垂直于导数的方向求平均可以平滑该结果，如例中所示。

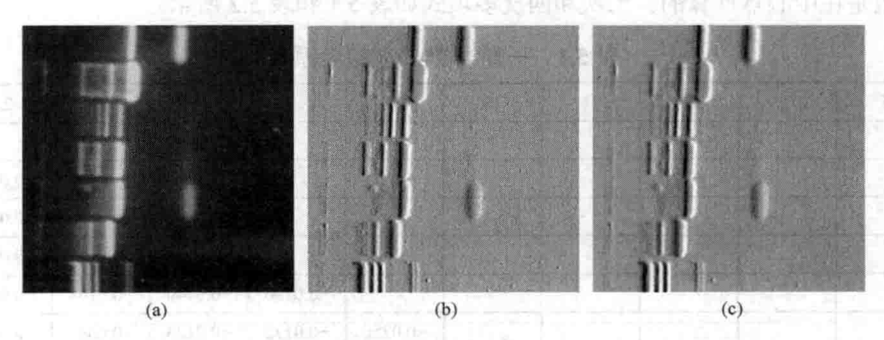


图 5.25　蛋白质分离凝胶图像：(a)原始图像；(b)使用 1 像素高的核的水平求导结果；(c)为降噪目的使用 5 像素高的核的水平求导结果

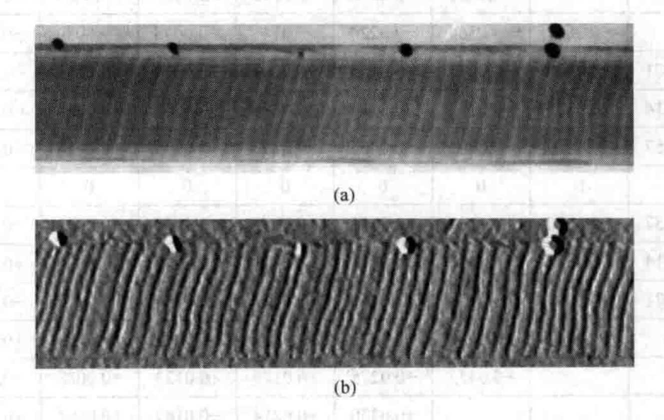


图 5.26　岩芯的年轮图像：(a)原始图像；(b)使用 3×3 核水平求导的结果

水平方向一阶求导的典型核是

$$\begin{bmatrix}+1 & \mathbf{0} & -1\end{bmatrix}\quad\begin{bmatrix}+1 & 0 & -1\\+2 & \mathbf{0} & -2\\+1 & 0 & -1\end{bmatrix}\quad\begin{bmatrix}+1 & 0 & -1\\+2 & 0 & -2\\+4 & \mathbf{0} & -4\\+2 & 0 & -2\\+1 & 0 & -1\end{bmatrix}$$

如同前面的章节中那样，值被计算的像素位置用粗体显示。这三个核间的区别是在垂直方向上执行的平均的数量。在所有这些核中，原始中心像素的值根本不参与计算。相反，是在左右邻点之间形成差。显然，这些模式可以旋转到其他方向，例如

$$\searrow \quad \begin{bmatrix} +1 & & \\ & \mathbf{0} & \\ & & -1 \end{bmatrix} \quad \begin{bmatrix} +2 & +1 & 0 \\ +1 & \mathbf{0} & -1 \\ 0 & -1 & -2 \end{bmatrix}$$

$$\downarrow \quad \begin{bmatrix} +1 \\ \mathbf{0} \\ -1 \end{bmatrix} \quad \begin{bmatrix} +1 & +2 & +1 \\ +1 & \mathbf{0} & -1 \\ -1 & -2 & -1 \end{bmatrix}$$

为看到正和负结果，结果中添加了值为 128（中间灰度）的偏移。

我们可以设计其他的核值，且通过使用实数而不是整数，可计算任何中间方向的导数。导数图像的外观通常描述为产生“浮雕”效果。由于它们基本上是一维导数，因此可以使用 Savitsky and Golay（1964）的系数，这些系数最初发布用于诸如光谱图一样的一维数据或其他纸带记录器输出。如同第 4 章中给出的平滑权重一样，这些系数相当于对数据的最小二乘高阶多项式拟合。但在拟合情形下，多项式的一阶导数是在中心点计算的。二次和四次多项式如表 5.1 和表 5.2 所示。

表 5.1　一阶导数二次拟合系数

#	5	7	9	11	13	15	17	19	21	23	25
−12											−0.0092
−11										−0.0109	−0.0085
−10									−0.0130	−0.0099	−0.0077
−9								−0.0158	−0.0117	−0.0089	−0.0069
−8							−0.0196	−0.0140	−0.0104	−0.0079	−0.0062
−7						−0.0250	−0.0172	−0.0123	−0.0091	−0.0069	−0.0054
−6					−0.0330	−0.0214	−0.0147	−0.0105	−0.0078	−0.0059	−0.0046
−5				−0.0455	−0.0275	−0.0179	−0.0123	−0.0088	−0.0065	−0.0049	−0.0038
−4			−0.0667	−0.0364	−0.0220	−0.0143	−0.0098	−0.0070	−0.0052	−0.0040	−0.0031
−3		−0.1071	−0.0500	−0.0273	−0.0165	−0.0107	−0.0074	−0.0053	−0.0039	−0.0030	−0.0023
−2	−0.2000	−0.0714	−0.0333	−0.0182	−0.0110	−0.0071	−0.0049	−0.0035	−0.0026	−0.0020	−0.0015
−1	−0.1000	−0.0357	−0.0250	−0.0091	−0.0055	−0.0036	−0.0025	−0.0018	−0.0013	−0.0010	−0.0008
0	0	0	0	0	0	0	0	0	0	0	0
+1	+0.1000	+0.0357	+0.0250	+0.0091	+0.0055	+0.0036	+0.0025	+0.0018	+0.0013	+0.0010	+0.0008
+2	+0.2000	+0.0714	+0.0333	+0.0182	+0.0110	+0.0071	+0.0049	+0.0035	+0.0026	+0.0020	+0.0015
+3		+0.1071	+0.0500	+0.0273	+0.0165	+0.0107	+0.0074	+0.0053	+0.0039	+0.0030	+0.0023
+4			+0.0667	+0.0364	+0.0220	+0.0143	+0.0098	+0.0070	+0.0052	+0.0040	+0.0031
+5				+0.0455	+0.0275	+0.0179	+0.0123	+0.0088	+0.0065	+0.0049	+0.0038
+6					+0.0330	+0.0214	+0.0147	+0.0105	+0.0078	+0.0059	+0.0046
+7						+0.0250	+0.0172	+0.0123	+0.0091	+0.0069	+0.0054
+8							+0.0196	+0.0140	+0.0104	+0.0079	+0.0062
+9								+0.0158	+0.0117	+0.0089	+0.0069
+10									+0.0130	+0.0099	+0.0077
+11										+0.0109	+0.0085
+12											+0.0092

表 5.2　一阶导数四次拟合系数

#	5	7	9	11	13	15	17	19	21	23	25
−12											+0.0174
−11										+0.0200	+0.0048
−10									+0.0231	+0.0041	−0.0048
−9								+0.0271	+0.0028	−0.0077	−0.0118
−8							+0.0322	−0.0003	−0.0119	−0.0159	−0.0165
−7						+0.0387	−0.0042	−0.0182	−0.0215	−0.0209	−0.0190
−6					+0.0472	−0.0123	−0.0276	−0.0292	−0.0267	−0.0231	−0.0197
−5				+0.0583	−0.0275	−0.0423	−0.0400	−0.0340	−0.0280	−0.0230	−0.0189
−4			+0.0724	−0.0571	−0.0657	−0.0549	−0.0431	−0.0335	−0.0262	−0.0208	−0.0166
−3		+0.0873	−0.1195	−0.1033	−0.0748	−0.0534	−0.0388	−0.0320	−0.0219	−0.0170	−0.0134
−2	+0.0833	−0.2659	−0.1625	−0.0977	−0.0620	−0.0414	−0.0289	−0.0210	−0.0157	−0.0120	−0.0094
−1	−0.6667	−0.2302	−0.1061	−0.0575	−0.0346	−0.0225	−0.0154	−0.0110	−0.0081	−0.0062	−0.0048
0	0	0	0	0	0	0	0	0	0	0	0
+1	+0.6667	+0.2302	+0.1061	+0.0575	+0.0346	+0.0225	+0.0154	+0.0110	+0.0081	+0.0062	+0.0048
+2	−0.0833	+0.2659	+0.1625	+0.0977	+0.0620	+0.0414	+0.0289	+0.0210	+0.0157	+0.0120	+0.0094
+3		−0.0873	+0.1195	+0.1033	+0.0748	+0.0534	+0.0388	+0.0320	+0.0219	+0.0170	+0.0134
+4			−0.0724	+0.0571	+0.0657	+0.0549	+0.0431	+0.0335	+0.0262	+0.0208	+0.0166
+5				−0.0583	+0.0275	+0.0423	+0.0400	+0.0340	+0.0280	+0.0230	+0.0189
+6					−0.0472	+0.0123	+0.0276	+0.0292	+0.0267	+0.0231	+0.0197
+7						−0.0387	+0.0042	+0.0182	+0.0215	+0.0209	+0.0190
+8							−0.0322	+0.0003	+0.0119	+0.0159	+0.0165
+9								−0.0271	−0.0028	+0.0077	+0.0118
+10									−0.0231	−0.0041	+0.0048
+11										−0.0200	−0.0048
+12											−0.0174

将求导的方向调整到垂直于一幅图像中的主要结构，可产生有用的结果，如图 5.27 所示。但要注意的是，尽管这种所谓的“浮雕”效果增强了在土星图片中云图案和部分环的能见度，但它也有效地隐藏了平行于求导方向的细节，而基于拉普拉斯算子的锐化则突出了该细节而与方向无关。

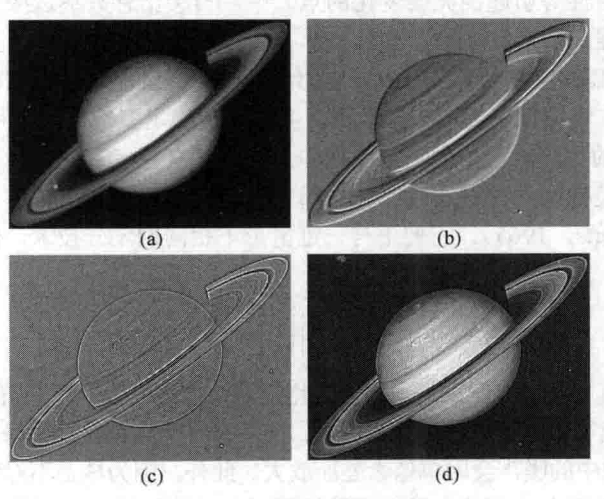


图 5.27　方向导数与拉普拉斯算子的比较：(a)原始图像；(b)从左上方应用求导的结果；(c)应用非方向拉普拉斯核的结果；(d)图(c)与图(a)相加后的结果

抑制某个特定方向的细节的能力是一种有用的性质，这种性质有时可用来揭示或增强图像中的其他信息。图 5.28 和图 5.29 说明了这种可能性。

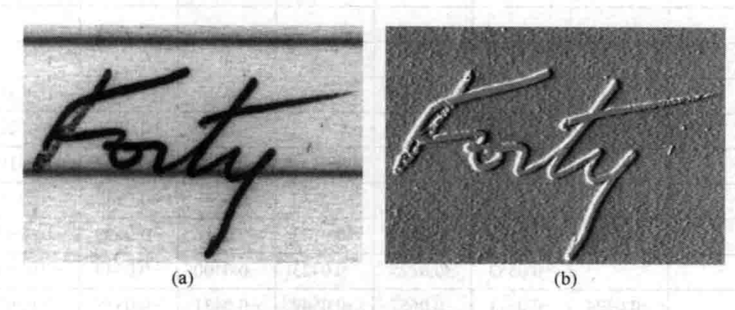


图 5.28　对笔迹图像(a)应用水平求导隐藏了纸上的格线(b)

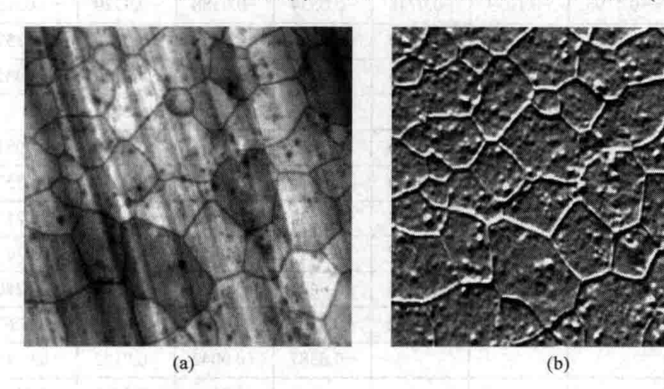


图 5.29　平行于铝箔图像的轧制方向应用求导，抑制了滚痕且使得金属晶粒结构更为清晰

5.6　使用梯度找到边缘

综上所述可知，拉普拉斯算子（二阶导数）并不是划定边缘的理想工具（Berzins，1984；Heath et al.，1997）。多数情形下，特征或区域的边界或边缘至少表现为亮度阶梯，且有时会跨域几个像素。拉普拉斯算子的适用特征，从有效性排序来看分别是点、线，尔后才是阶。在带噪图像（噪声通常表现为由计数统计、检测器特性等引起的亮度变化的点）中，与突出感兴趣边缘或边界相比，拉普拉斯运算突出这些点的能力更强。

方向一阶导数仅突出那些垂直于它们的方向的边缘，而通过沿边缘平均可降低它们对噪声的敏感性。但使用方向导数从二维图像中提取一维数据，并不是能满足多数真实图片需求的专用操作。

将方向导数中所用的相同原理用到二维图像中来突出具有任意方向的边界，是最常用的图像增强操作之一。当然，问题是找到一种对边缘的（局部）方向不敏感的方法。这一任务的最早方法之一是 Roberts 交叉算子（Roberts，1965）。它使用与一维情形下相同的差分技术，但彼此间有两个像素差，所用的核为

$$\begin{bmatrix} +1 & 0 \\ 0 & -\mathbf{1} \end{bmatrix} \begin{bmatrix} 0 & +1 \\ -1 & \mathbf{0} \end{bmatrix}$$

这两个差代表对亮度导数的有限逼近。组合使用这两个方向导数，可获得对边缘方向不敏感的幅度值，方法是先平方、再相加，尔后取总和的平方根。与用于一个方向上的差分方法相比，这种方法也存在许多相同的问题。图像中的噪声会因单像素差而放大。此外，因为核是不对称的，结果在 x 和 y 方向上都会偏移半个像素。事实上，在首次提出这种模型时，计算机速度并不快，尤其是求平方根的浮点数学运算。因此，还用到了两种其他方法：将两个方向差的绝对值相加，或比较两个差的绝对值而保

留较大者。这两种方法都可在某种程度上使得结果对方向敏感。

如同上面的水平导数例子，使用更大的核可降低对噪声的敏感性，方法是平均几个像素并消除图像偏移。3×3 的求导核得到了广泛使用。常用系数的几个例子如表 5.3 所示。

表 5.3　方向导数滤波器权重的几个例子

A	B	C	D
$\begin{bmatrix}+1 & 0 & -1\\+1 & \mathbf{0} & -1\\+1 & 0 & -1\end{bmatrix}$	$\begin{bmatrix}+1 & 0 & -1\\+2 & \mathbf{0} & -2\\+1 & 0 & -1\end{bmatrix}$	$\begin{bmatrix}+1 & -1 & -1\\+2 & \mathbf{+1} & -1\\+1 & -1 & -1\end{bmatrix}$	$\begin{bmatrix}+5 & -3 & -3\\+5 & \mathbf{0} & -3\\+5 & -3 & -3\end{bmatrix}$
$\begin{bmatrix}+1 & +1 & 0\\+1 & \mathbf{0} & -1\\0 & -1 & -1\end{bmatrix}$	$\begin{bmatrix}+2 & +1 & 0\\+1 & \mathbf{0} & -1\\-1 & -1 & -2\end{bmatrix}$	$\begin{bmatrix}+2 & +1 & -1\\+1 & \mathbf{+1} & -1\\-1 & -1 & -1\end{bmatrix}$	$\begin{bmatrix}+5 & +5 & -3\\+5 & \mathbf{0} & -3\\-3 & -3 & -3\end{bmatrix}$
$\begin{bmatrix}+1 & +1 & +1\\0 & \mathbf{0} & 0\\1 & -1 & -1\end{bmatrix}$	$\begin{bmatrix}+1 & +2 & +1\\0 & \mathbf{0} & 0\\-1 & -2 & -1\end{bmatrix}$	$\begin{bmatrix}+1 & +2 & +1\\-1 & \mathbf{+1} & -1\\-1 & -1 & -1\end{bmatrix}$	$\begin{bmatrix}+5 & +5 & +5\\-3 & \mathbf{0} & -3\\-3 & -3 & -3\end{bmatrix}$

只要结果的大小不超过所用软件的存储容量，那么使用哪种值的模式几乎没有区别。若这被限制到单个字节每像素，则结果必须在-127 和+128 之间，因为上面运算的结果要么为负，要么为正。大亮度阶可能会截断计算的值，进而为实现看到较小边界的目的而加宽或扭曲主要的边界。可供替代的选择是采用自动标定方法，即使用导数图像中的最大值和最小值来设置白值和黑值。为了避免精度损失，这需要遍历图像两次：第一次是进行计算并找到极值，第二次是为存储目的标定结果。

对于 Roberts 交叉算子法，如果计算了两个正交方向的导数，那么可通过计算它们的平方和的平方根，来得到与方向无关的结果。这是计算水平方向和垂直方向的两个矢量和，来得到表示亮度变化局部梯度的矢矩的常用方法。

$$\text{幅度} = \sqrt{\left(\frac{\partial B}{\partial x}\right)^2 + \left(\frac{\partial B}{\partial y}\right)^2} \tag{5.6}$$

这就是 Sobel（1970）方法。它是最常用的技术之一（至于 Roberts 交叉算子，有些早期的计算机程序试图通过相加或比较两个值而非求平方、相加并取平方根来弥补硬件的局限）。图 5.30 显示了方向导数及它们的组合。

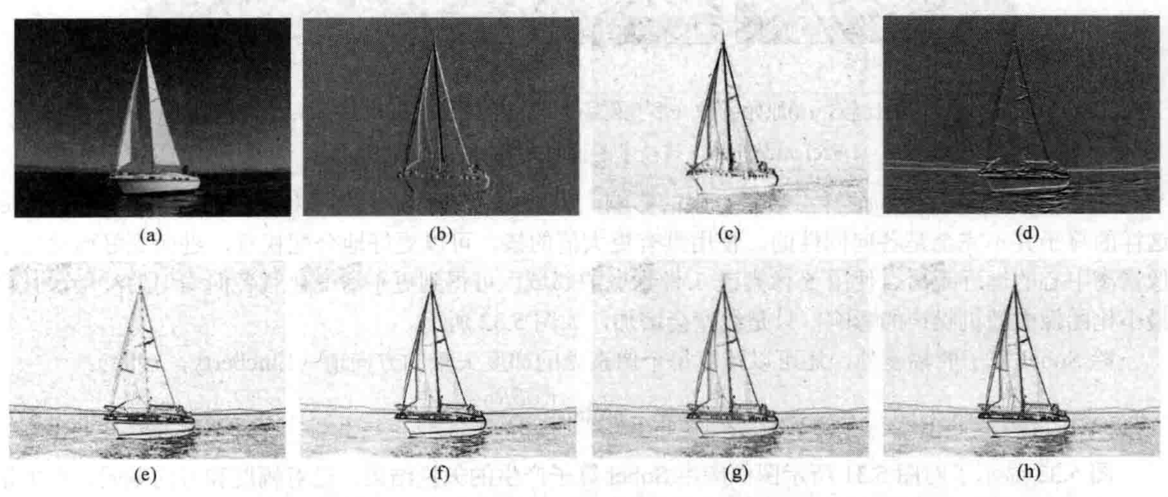


图 5.30　方向导数和 Sobel 边缘算子：(a)原始图像；(b)水平导数（“浮雕图像”）；(c)水平导数的绝对值；(d)垂直导数（“浮雕图像”）；(e)垂直导数的绝对值；(f)通过保留较大者来合并导数绝对值；(g)通过相加来合并导数的绝对值；(h)通过求平方、相加并取平方根来合并导数（Sobel 算子）

采用适当的硬件，如移位寄存器或阵列处理器，基本上可实时执行 Sobel 运算。也就是说，要处理和查看传统的视频图像，对每图像的计算时间平均为 1/30 秒或 1/60 秒。有些处理运算要求在查看一帧图像的同时，数字化并处理下一帧图像，但在 Sobel 情形下则可在仅延迟两条视频扫描线的情形下实时观看图像。两条数据线可被缓冲并用于计算水平和垂直导数值（使用上面给出的 3×3 核）。某些军事应用中会使用专用硬件来执行这一实时边缘增强，如在飞机空中加油时，跟踪和对齐图像中的边缘。

Kirsch 算子（Kirsch，1971）避免了 Sobel 中需要去计算平方和的平方根的数学运算。这种方法应用导数核的八个方向，并保留最大值。它只需要整数乘法与比较运算。对于许多图像，边缘的幅度与 Sobel 的非常类似。

图 5.31 比较了拉普拉斯算子、Roberts 交叉算子与 Kirsch 算子。示例图像中包含了碳膜中围绕孔洞各个方向的连续边缘，以及沿石棉纤维各个方向的直边。拉普拉斯图像中的噪声很多，Roberts 交叉算子同样地不能很好地显示所有边。对于这幅图像而言，Kirsch 算子的结果在视觉上与 Sobel 的相似。

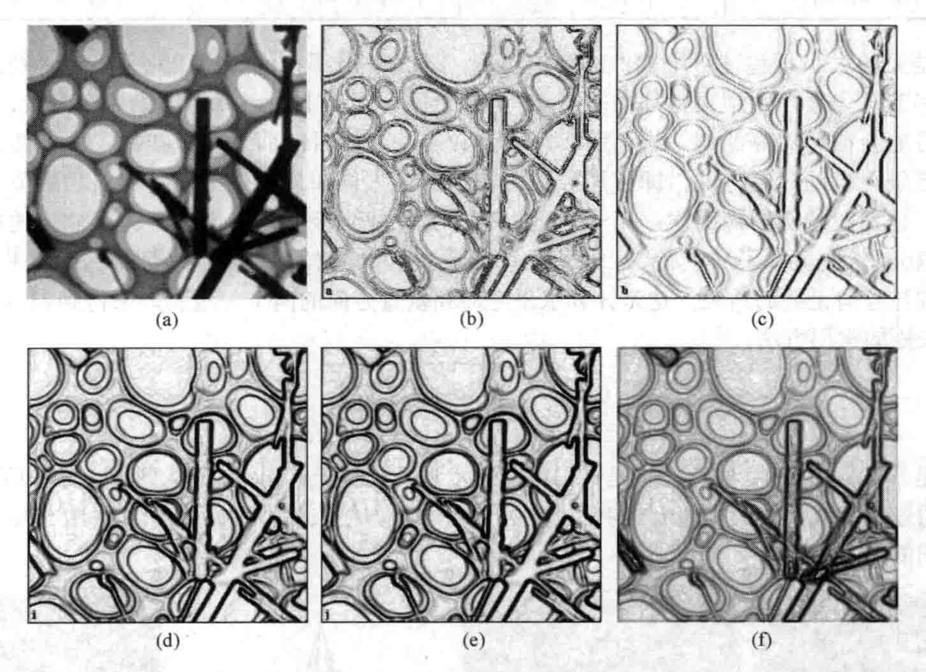


图 5.31　边缘描述算子的比较：(a)原始图像（多孔碳膜上的石棉纤维，TEM 成像）；(b)拉普拉斯；(c)Robert 交叉；(d)Sobel；(f)Frei and Chen（该技术的说明见下方，在此处引入这一结果的目的是便于对比）

由于方形像素网格和抽样点数量有限的影响，以及仅使用局部像素值对噪声的敏感性，像 Sobel 这样的算子并不完全是各向同性的。使用具有更大值的核，可以更好地分配权重，进而更好地对应于像素离中心的远近距离。使用 5 像素或 7 像素宽的邻域，可得到更平滑变化且各向均匀的结果，且可最小化图像中随机噪声的影响，只是线宽会增加，如图 5.32 所示。

除 Sobel 算子的幅度外，还可以计算每个像素处的梯度矢量的方向值（Lineberry，1982）：

$$方向 = \arctan\left(\frac{\partial B/\partial y}{\partial B/\partial x}\right) \tag{5.7}$$

图 5.33 显示了对图 5.31 所示图像应用 Sobel 算子产生的矢量结果。已对幅度和方向编码，但矢量场太稀疏以至于不能显示任何图像细节。如果该计算用于将一个值赋给梯度方向的每个像素，那些这些值可标定到图像的灰度。如果使用从 0 到 255 的典型像素值来表示从 0° 到 359° 的角度，那么灰度中的每个阶对应于约 1.4° 的矢量方向。

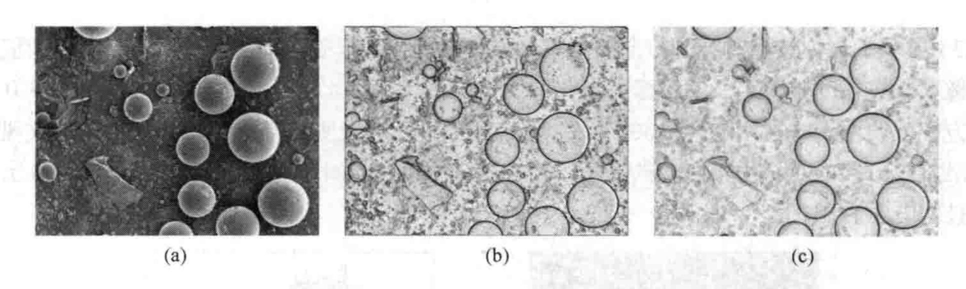


图 5.32 邻域大小对边缘描述的影响：(a)原始图像（球形颗粒的 SEM 图像）；(b)3 像素宽的邻域；(c)7 像素宽的邻域

图 5.33(c)中各个接近圆孔周围的值的变化是明显的。为显示目的使用伪彩色标度特别适合，因为色调的彩虹式标度能显示灰度所需要的在零度处无不连续的这种变化。在彩色图像［见图 5.33(d)］中，结合 Sobel 梯度矢量的幅度和角度，能更清晰地显示边缘的位置和方向。在每个点处，向量的幅度被赋为亮度通道，而角度被赋为色相通道，饱和度则设置到最大。

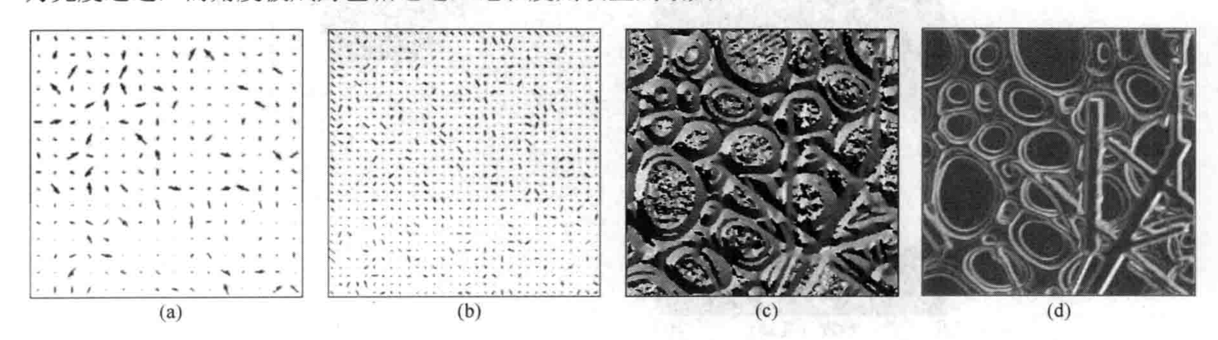


图 5.33 对图 5.31(a)所示图像应用一个 Sobel 算子得到的局部梯度方向：(a)一个粗矢量场，其中的每个矢量均有由 Sobel 算子给出的方向和幅度；(b)细网格场，它仍不足以显示结构；(c)对每个像素的亮度赋一个矢量角；(d)由幅度（灰度）和方向（色样）组合表示的颜色

也可以对矢量幅度设置一个阈值，并仅为高于该阈值的那些像素保存角度信息。在图 5.34 中，这样做既可对边缘像素分配颜色，也可保存角度值的直方图，进而确定边缘的方向。

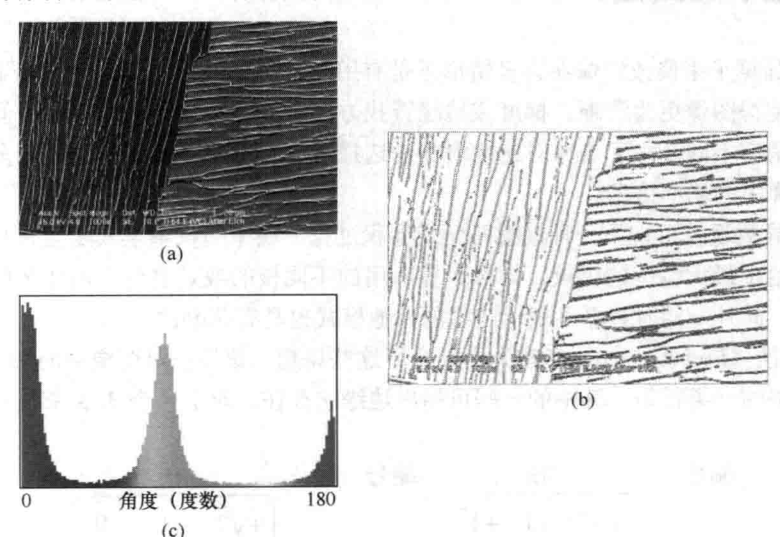


图 5.34 钛合金中的薄层状结构：(a)SEM 图像；(b)由 Sobel 梯度的幅度和方向编码的颜色来取的边缘；(c)从 0° 到 180° 的边缘方向的直方图

图 5.35 显示了这种技术对纤维方向的测量结果。由于纤维覆盖整幅图像，因此不需要阈值来选择将被测量的像素。按亮度梯度定义的边缘方向与纤维反面的像素相比，有 180° 的不同。覆盖从 0° 到 360° 范围的直方图表明两个峰之间间隔 180°（128 个亮度值），即表明纤维不是各向同性的。玫瑰图表明，两倍以上沿纤维方向的点与水平方向的夹角约为 70°。应用 Sobel 算子的这种效率以及可绘出直方图，使得这种技术非常有用。

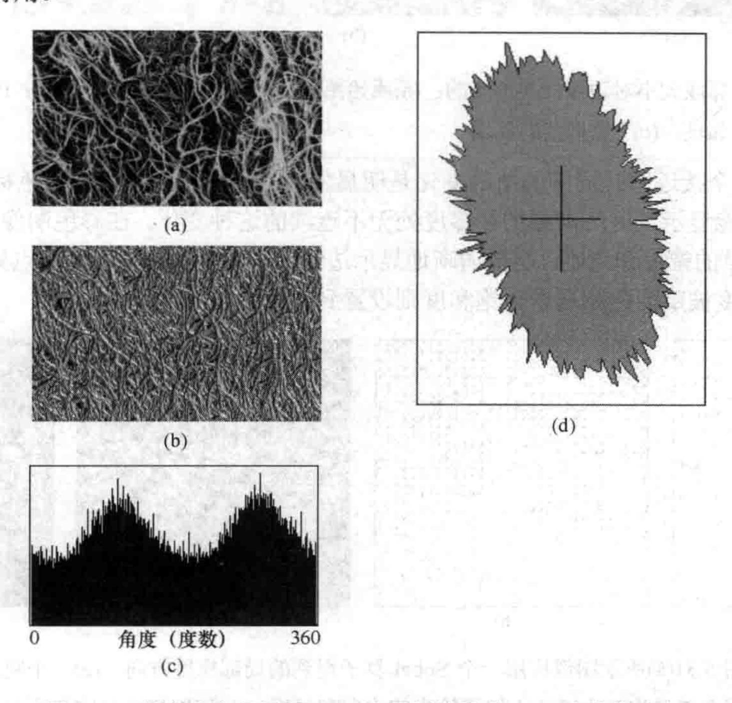


图 5.35　纤维方向的测量：(a)胶原纤维图像；(b)赋给像素亮度的 Sobel 方向值；(c)图(b)中从 0° 到 360° 的直方图

5.7　其他边缘检测器

使用边缘增强算子来修改图像在许多情形下是有用的。例如，使用拉普拉斯算子进行锐化可增大边缘的对比度，使得图像更为清晰。梯度或边缘查找方法也可以达到这样的效果，但也修改图像，因此其解释会变得稍有不同。由于这种对比度增加是选择性的，因此它以一种不控制全部亮度直方图的方式来对应于图像中的局部信息。

使用核进行导数运算可视为一种模板匹配或卷积过程。核中的权重模式是当它与图像像素的亮度值模式匹配时，给出最大响应的模板。导数运算所用的不同核的数量表明，由什么构成边界并不存在单一的最佳定义。此外，同时查看不表示边缘的其他模式也是有帮助的。

Frei-Chen 算法（Frei & Chen，1977）中结合了这些思想，该算法对图像中的每个点应用一组核。每个核提取图像中的一类行为，其中的一些可指出边缘的存在。对于一个 3×3 邻域，作为独立基函数的核显示如下。

编号	核	编号	核
0	$\begin{bmatrix} +1 & +1 & +1 \\ +1 & \mathbf{+1} & +1 \\ +1 & +1 & +1 \end{bmatrix}$	4	$\begin{bmatrix} +\sqrt{2} & -1 & 0 \\ -1 & \mathbf{0} & +1 \\ 0 & +1 & -\sqrt{2} \end{bmatrix}$

$$1\quad\begin{bmatrix}-1 & -\sqrt{2} & -1\\ 0 & \mathbf{0} & 0\\ +1 & +\sqrt{2} & +1\end{bmatrix}\qquad 5\quad\begin{bmatrix}0 & +1 & 0\\ -1 & \mathbf{0} & +1\\ 0 & -1 & 0\end{bmatrix}$$

$$2\quad\begin{bmatrix}-1 & 0 & +1\\ -\sqrt{2} & \mathbf{0} & +\sqrt{2}\\ -1 & 0 & +1\end{bmatrix}\qquad 6\quad\begin{bmatrix}-1 & 0 & +1\\ 0 & \mathbf{0} & 0\\ +1 & 0 & -1\end{bmatrix}$$

$$3\quad\begin{bmatrix}0 & -1 & +\sqrt{2}\\ +1 & \mathbf{0} & -1\\ -\sqrt{2} & +1 & 0\end{bmatrix}\qquad 7\quad\begin{bmatrix}+1 & -2 & +1\\ -2 & \mathbf{+4} & -2\\ +1 & -2 & +1\end{bmatrix}$$

$$8\quad\begin{bmatrix}-2 & +1 & -2\\ +1 & \mathbf{+4} & +1\\ -2 & +1 & -2\end{bmatrix}$$

只有核 1 和核 2（水平和垂直梯度检测器）能指出边缘的存在。因此，对每个像素应用每个核，然后将结果相加，可生成核 1 和核 2 的结果与其他核的结果之比。该比率的平方根的余弦被用做“边缘状态”的度量，并赋给导数图像中的像素位置。

与诸如 Sobel 这种更为传统的边缘检测器相比，优势是对相对像素值的配置更灵敏，而与因在图像中的位置不同而不同的亮度幅度无关。图 5.31(f)和图 5.36(a)显示了 Frei-Chen 算子的例子，它们可以与其他几种技术进行比较。

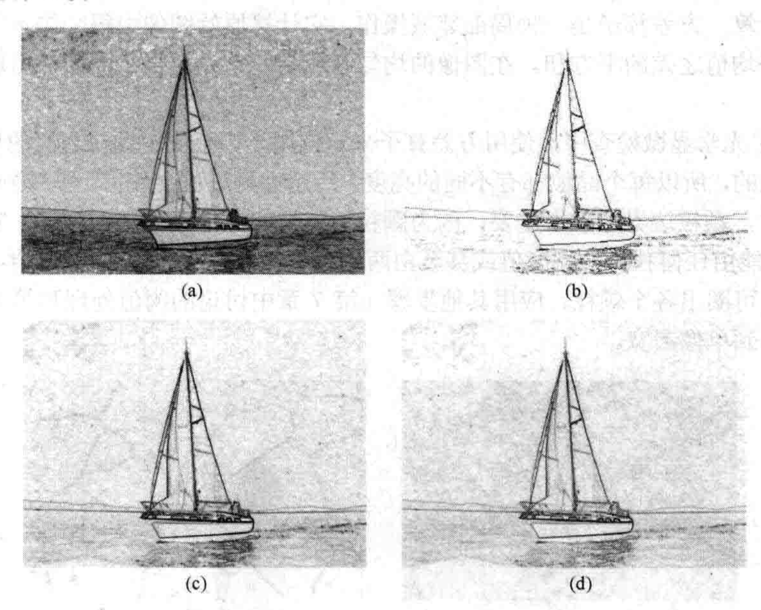


图 5.36　对图 5.30(a)所示图像应用几种边缘查找方法的比较：(a)Frei and Chen 法；(b)方差法；(c)最大一最小法［方法(c)和(d)的介绍见下方，此处包含它们的目的在于方便比较］

前面所有的边缘查找方法，都会生成沿边缘的线，这些线较宽，且只近似地定位了实际边界。Canny 滤波器（Canny，1986；Olsson，1993）试图将边缘定位到最可能的单像素位置［见图 5.36(b)］。这是一个强大的多步函数，它使用了上面描述并演示的几种技术。它首先对图像进行高斯平滑来降噪，然后使用较大的核来计算两个正交方向的导数，并将导数结果合并来得到一个矢量幅度和方向。接着使用非最大抑制技术来使沿边缘和不连续点处的线条变细。将每个点处的梯度幅度与梯度方向中下一像素的值进行比较，保留其中的较大者。这是一种可应用到任何边缘查找算子来减小线宽的技术（Bertrand et al.，1997；Neubeck & VanGool，2006；Guo et al.，2010；Sun & Vallotton，2010）。最后，仅保留那些梯度值超过某

个预设阈值的线条。由于 Canny 法有几个用户定义的任意常数，因此不具有通用性，但在某些机器视觉应用中，可用这一技术来精确地定位边缘。图 5.36 和图 5.37 比较了 Canny 与其他边缘检测器的结果。

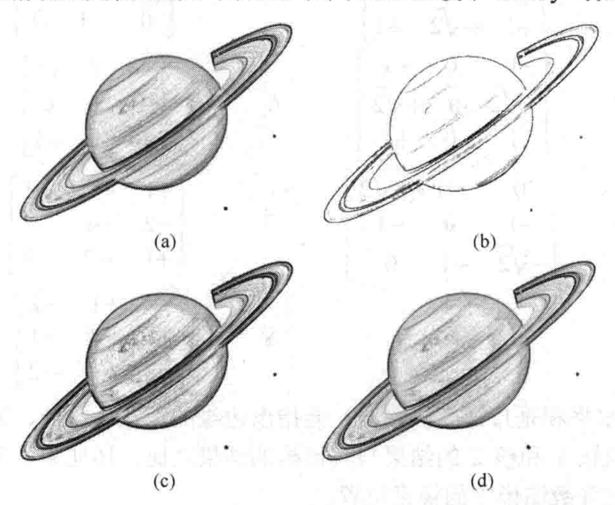


图 5.37　对图 5.27(a)所示图像应用几种边缘查找方法的比较：(a)Sobel 法；(b)Canny 法；(c)方差法；(d)最大一最小法

图 5.36(c)和图 5.37(c)显示了使用一种不同方法的几个例子，这种方法并不基于与权重核的卷积，而基于一种统计计算。方差算子是一种局部邻域操作，它计算原始图像中每个像素的亮度值与该像素周围邻域的亮度平均值之差的平方和。在图像的均匀区域中，该方差很小，而在出现台阶或线条时，它会变得很大。

图 5.38 显示了光学显微检查中，使用方差算子来定位抛光铝金属中颗粒之间的边界。因为晶格方向在空间中是随机的，所以每个晶粒都有不同的亮度，所用蚀刻过程会使得一些颗粒比其他颗粒更暗。在研究金属结构时，颗粒边界通常很重要，因为颗粒边界形状源于热处理和加工，它会影响到许多机械性能。边界并不能由任何独特的亮度值或甚至由两侧值的任何一致差值业识别。半径为 2.5 像素的圆形邻域的方差值，可圈出各个颗粒。应用其他步骤（第 7 章中讨论的阈值处理和第 8 章中介绍的骨架化）可将轮廓减小到单像素宽。

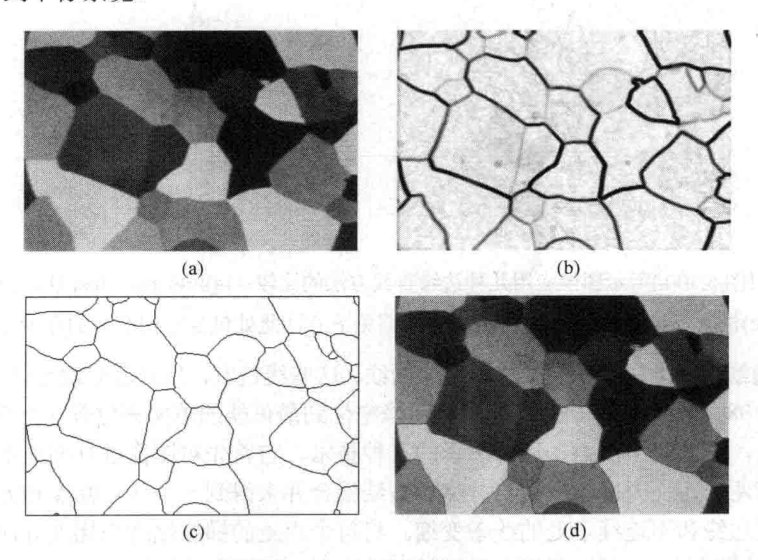


图 5.38　颗粒间边界的划定：(a)铝金属，经抛光和腐蚀以显示不同的颗粒；(b)方差边缘查找算法的结果；(c)阈值处理和骨架化会在边界处产生 1 像素宽的单线；(d)叠加到原始图像上的边界

人们对边缘查找算子感兴趣的原因有多种。一是边界的位置有助于各种测量操作；二是边缘可形成场景的"原始简图"，进而将这种简图用于高级软件中，如在机器人的软件中，根据所出现物体的几何形状来理解场景。艺术家或漫画家会使用边缘来表达对物体和结构的感知。如图 5.39 所示，边缘查找过程可完成类似的任务。

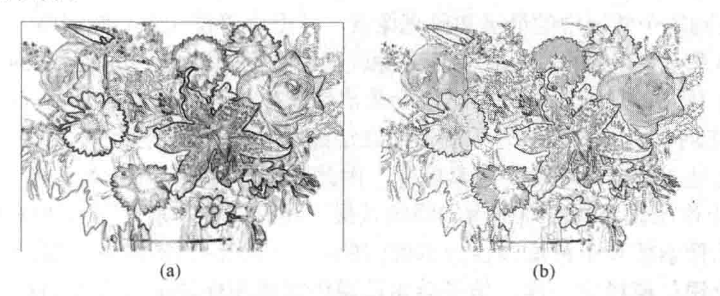


图 5.39　使用边缘算子将图像转换为素描［原始图像如图 5.6(a)所示］：(a)分别对红色、绿色和蓝色通道应用 Sobel 算子；(b)对 L 通道应用 Canny 算子

另一个目的是方便图像的后续处理。如上所述，图像锐化是通过增大边缘处的对比度来进行的。但该过程也会增大图像中远离边缘区域中的随机噪声的可见性。使用基于边缘查找算子结果的权重，可混合锐化算子的结果与平滑算子的结果，在锐化边缘的同时，降低远离边缘区域中的噪声，改善图像的整体外观（Kotera & Wang，2005；Kim & Allebach，2005）。数码摄影师经常会使用到这种方法，一些消费型相机的固件中，甚至集成了这种方法。图 5.40 给出了这样一个例子。这是自适应或条件处理方法的一个例子，它依据每个像素周围的局部邻域来应用不同的滤波器（在这种情形下，有几种可能性）。

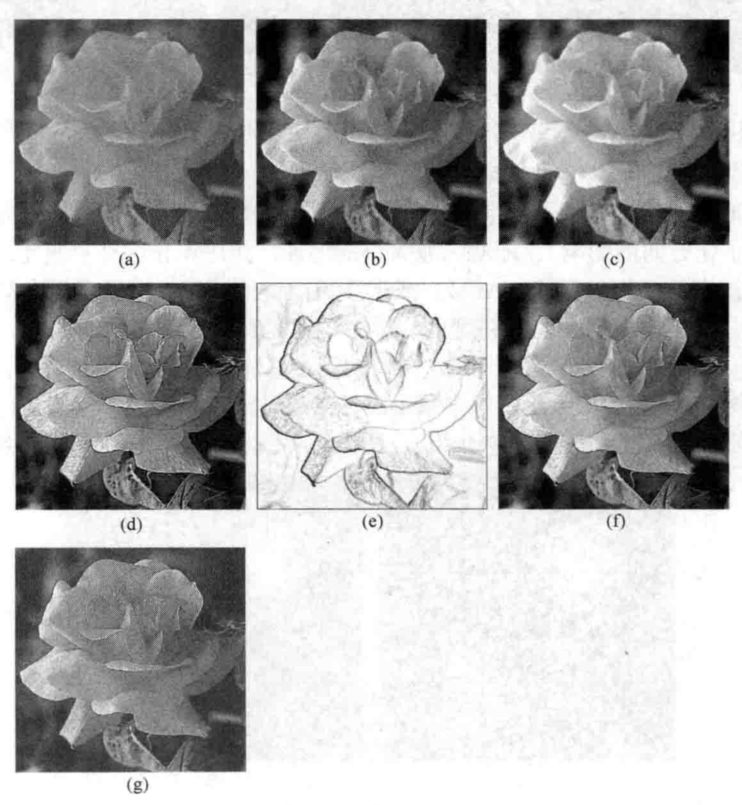


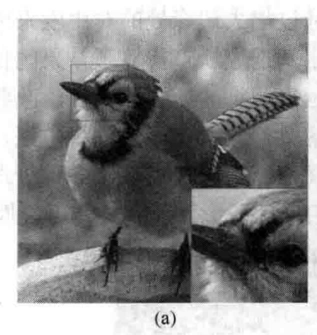
图 5.40　锐化边缘同时平滑均匀区域的条件处理示例：(a)原始图像；(b)亮度通道；(c)对图(b)应用高斯平滑的结果；(d)对图(b)应用拉普拉斯锐化的结果；(e)对图(b)应用边缘检测的结果，结果的幅度用做权值，以进一步混合平滑和锐化后的结果

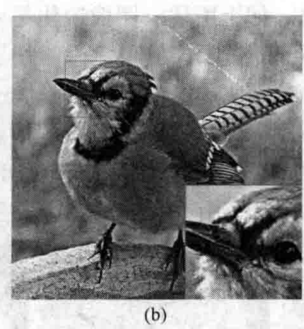
前一节中讨论的多数邻域运算会使用算术运算（相乘、相加等）来组合各个像素的值。另一类的算子也会使用到邻域，但这时的操作是比较和排序。有些此类操作类似于上面描述的方法。

第 4 章介绍的中值滤波器会按亮度对某个区域中的像素排序，找到中值，并用该值替换中心像素值。从图像中消除噪声的这一操作，会从图像中删除极端值。排序运算还包括最大和最小算子或膨胀和腐蚀算子，这类算子会找到每个邻域中的最暗和最亮像素，并用该值替代中心像素值。这些运算的一种用途如第 4 章中的图 4.50 所示，其目的是从图像中删除较细的特征，进而生成调匀后的背景。

一个简单的基于排序的边缘查找滤波器，会求出邻域中最亮和最暗像素值间的差。如图 5.36(d)和图 5.37(d)中所示，在均匀区域内，该差很小，而在边缘和台阶处该差会增大。这个最大－最小算子与方差算子有些类似之处，但由于对噪声像素敏感，因此通常不能很好地定位边缘。

第 4 章比较了中值滤波器与高斯平滑的降噪效果。还有一种基于排序的边缘增强技术，它可与反锐化掩模相比。原始像素值和中值滤波器计算值间的差，可加回到图像中，就像原始像素与高斯平滑结果间的差被加到反锐化掩模中一样。结果会如反锐化掩模那样增加边缘处的对比度，但还有几个重要的优点：在边缘周围不会引入光晕，边缘附近的小细节会得以保留，整个亮度范围内的对比度会更均匀。图 5.41 比较了这两个过程。两者都只应用到亮度通道，而颜色不变。


(a)


(b)

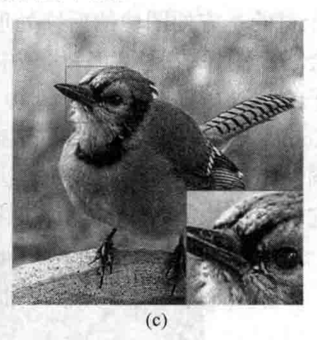
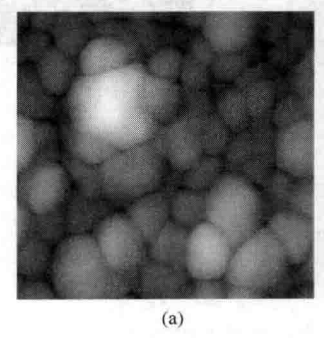

(c)

图 5.41　使用排序算子锐化：(a)原始图像；(b)反锐化掩模结果，高斯标准差为 3 像素；(c)基于排序锐化的结果，中值半径为 3 像素

使用排序算子时的重要变量之一是邻域的大小。通常，所用邻域的形态是方形（为了方便计算）或近似为圆形（最小化方向的影响）。随着邻域大小的增加，排序操作的计算量迅速增大。

距离图像（如源自原子力显微镜（AFM）或干涉显微镜的图像，或源自激光测距或雷达成像的图像）会为每个像素分配一个表示高程的亮度值。排序运算尤其适合于这样的图像，且通常用于定位边界。图 5.42 给出了一个淀积涂层地形的 AFM 图像的例子。先进行灰度腐蚀（使用 5 像素宽邻域中的最亮像素替代每个像素），然后从原始图像中减去腐蚀结果，沿边界生成一组线。再后进一步进行阈值处理和骨架化（分别在第 7 章和第 8 章中介绍），所得边界线将有助于涂层中各个结构的计数和测量。


(a)

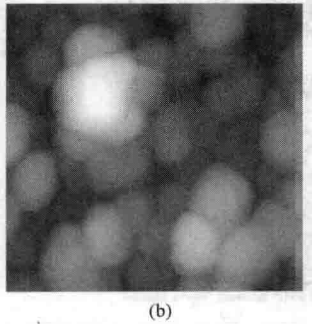

(b)

图 5.42　使用排序算子定位边缘：(a)原始图像（淀积涂层的 AFM 图像）；(b)使用排序滤波器，将每个像素替换为其最亮邻域像素；(c)图(b)和图(c)之差显示了清晰的边界；(d)对图像(c)进行阈值和骨架化处理得到边界线

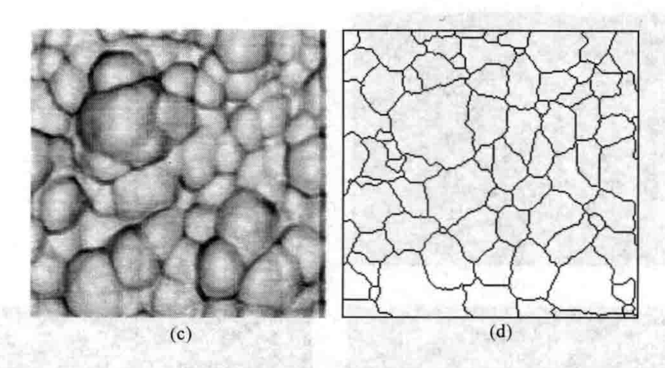


图 5.42 使用排序算子定位边缘：(a)原始图像（淀积涂层的 AFM 图像）；(b)使用排序滤波器，将每个像素替换为其最亮邻域像素；(c)图(b)和图(c)之差显示了清晰的边界；(d)对图像(c)进行阈值和骨架化处理得到边界线（续）

5.8 纹理

许多图像中包含有不由特殊亮度或颜色值表征，而由称为纹理的亮度值模式表征的区域。这一术语的定义有点松散，指的是从一个点到下一个点或在一个小区域内的局部亮度（或颜色）变化。在图像表示为表面时，如果亮度解释为高度，那么纹理则是表面粗糙度的度量，表面粗糙度是具有多重意义和测度的一个术语，详见第 15 章。

排序运算是可以用来检测图像中这种纹理的一种工具。上面介绍的边缘查找算子，是邻域中最大和最小亮度值间的差。对于平滑、平坦或均匀的区域，这一差值很小。较大的差值对应于具有较大感知纹理或粗糙度的区域。邻域尺寸必须大到足以包含暗和亮像素，即邻域尺寸要大于任何可能出现的小均匀细节。

图 5.43 显示了一个例子，其中原始图像的直方图有一个宽广的单峰，基于亮度我们在视觉上并不能区分平滑和有纹理的区域。最大－最小图像（使用 7 像素宽的圆形邻域）产生了允许阈值处理的不同亮度值。平滑区域（奶酪显微图像中的凝乳）产生了一个低的最大－最小值，而高度纹理化的蛋白质产生了一个更大的值。因为该操作是局部邻域处理，因此它也消除了图像中背靠背的阴影。

图中显示的第二种纹理提取方法是，在一个移动邻域（7 像素宽的圆形邻域）中计算像素值的统计方差，这种方法也能区分在图像中的区域。注意，采用一个较小的邻域尺寸，最大－最小算子和方差算子也都可用于定位边缘。为便于比较，图中还显示了下面将要说明的分形纹理算法的结果。

卫星图像通常可由纹理算子来计算。农作物、建筑物和其他土地利用可产生人类能够识别的清晰纹理。在软件中实现这种能力的算法已被人们找到。在关于该主题的一篇经典论文中，Haralick 列出了 14 个利用区域内像素及其亮度差此类纹理算子（Haralick et al.，1973；Haralick，1979；Weszka et al.，1976；Jiang et al.，2008）。所用区域不是围绕每个像素的邻域，而是包含由某些边界或其他标识规则（如亮度等）所圈出的连续块内所有像素的区域。他构建了一个区域内相邻像素对数量与亮度关系的表，然后使用该像素表来计算纹理参数。

在式(5.8)中，数组 $P(i,j)$中包含的是亮度分别 i 和 j 的最近邻像素对（90° 方向）的数量。R 是一个重新归一化常数，等于图像中或任何用于计算的矩形部分中，像素对的总数。原理上，它可扩展到由距离 d 分隔的像素对，或扩展到以 45° 方向对齐的像素对（其间距比 90° 方向的间距大）。对该区域中的所有像素对求和。Haralick 将此应用到了矩形区域，但它同样适用于不规则轮廓内的像素。

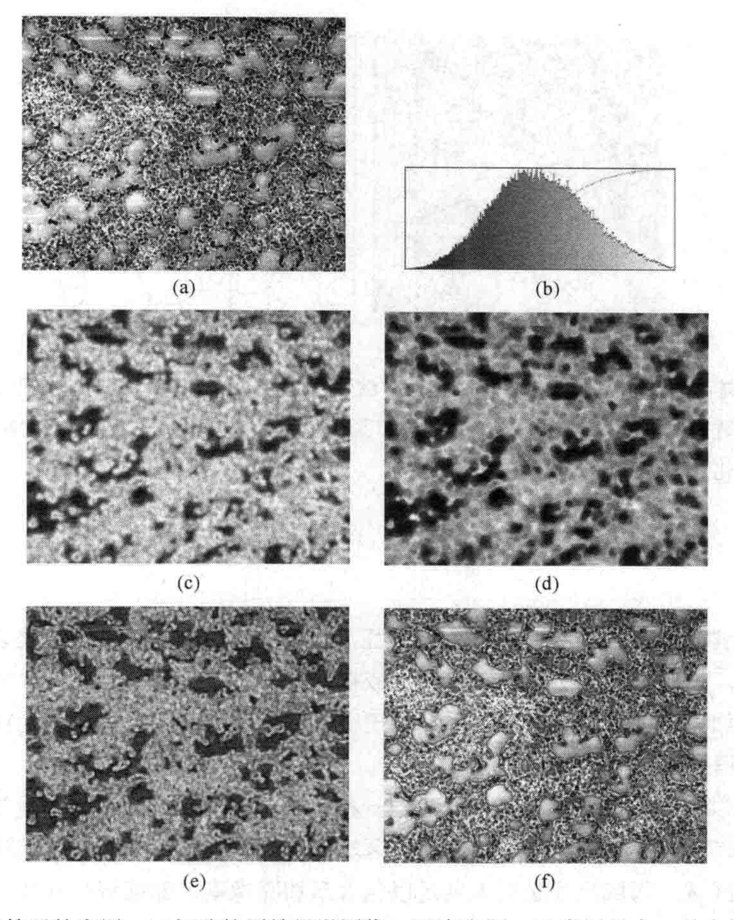


图 5.43　纹理算子的应用：(a)奶酪的原始显微图像；(b)直方图；(c)应用最大－最小算子后的结果（7 像素直径圆形邻域）；(d)应用方差算子后的结果（7 像素直径圆形邻域）；(e)分形纹理算子的结果（4 像素最大半径）；(f)图像(d)经阈值处理后，叠加到原始图像后的结果

所示的第一个参数是一个使用二阶矩的均匀性测度。由于各项被平方，因此几个较大并的贡献要大于许多较小差的贡献。所示的第二个参数是一个差值矩，它是图像中对比度的测度。第三个参数是图像中亮度线性相关性的测度，是通过相关运算得到的。

$$f_1=\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{P(i,j)}{R}\right)^2,\ f_2=\sum_{n=0}^{N-1}n^2\left\{\sum_{|i-j|=n}\left(\frac{P(i,j)}{R}\right)\right\},\ f_3=\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{ijP(i,j)}{R}\right)-\mu_x\mu_y}{\sigma_x\sigma_y} \tag{5.8}$$

式中，N 是灰度级的数量，μ 和 σ 分别是 x 和 y 方向上累积亮度值分布的均值和标准差。其他几个参数描述方差、熵和亮度值相关的信息测量。Haralick 已证明，当应用到卫星照片上的大矩形区域时，这些参数可以从草地中分辨出水体、分辨出不同的砂岩，从湿地或城区中分辨出林地，还可以分辨出不同的作物种植模式。

当在以每个像素为中心的一个移动邻域内计算时，可标定相同的值来创建导数图像，导数图像中，亮度差表示原始的纹理变化。在医学成像（Zizzari，2004）、显微学、遥感及其他方面，也应用了这些工具。在任何给出的情形下，可能需要试验几个纹理算子，以找到可最好分辨感兴趣特征与周围环境的那一个算子。图 5.44 给出了 Haralick 角二阶矩算子（上面的 f_2）的用法：将该算子应用到中心为每个像素的移动邻域来计算出一个纹理值，然后将纹理值分配给该像素。为便于比较，图中还显示了在一个圆形邻域（半径为 3）中计算出的熵以及后面将要说明的分形纹理计算的结果。

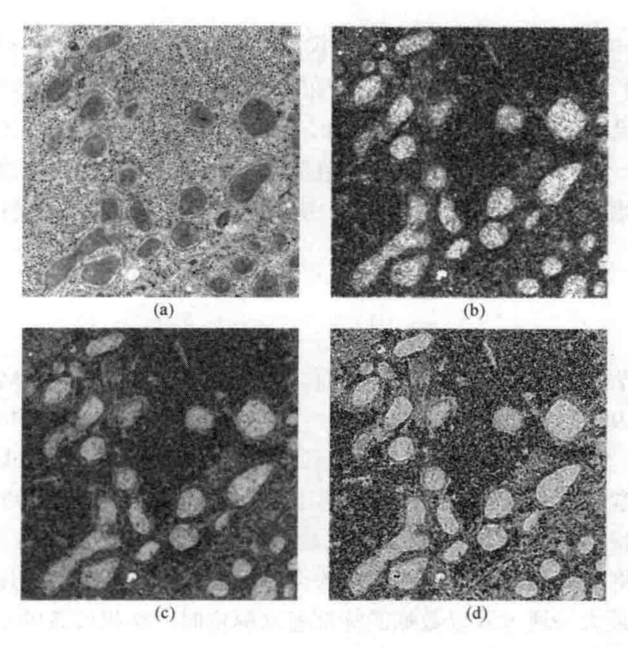


图 5.44　纹理算子的应用：(a)肝组织的原始 TEM 图像；(b)Haralick 角二阶矩；(c)熵；(d)最大半径为 6 像素的分形纹理计算结果

大范围统计、频率和像素变化的其他局部测度，被用于从图像中提取纹理的不同方面，如用于测量，或辅助将原始图像分割成各个区域（Heeger & Bergen，1995；Barth et al.，1998；de Bonet & Viola，1998；Ursani et al.，2008；Bergman et al.，2008）。这并不令人奇怪，因为所遇到的纹理类型多种多样，有些是规则的，有些则是随机的。不存在为各种情形选择最合适参数的方法。

结构和图像中出现的另一类纹理是有方向性的，或方向不是随机的。上面介绍的 Sobel 方向算子通常用于将局部亮度梯度的不同方向转换为有助于分割物体或结构的亮度差。图 5.45 显示了一个例子。人

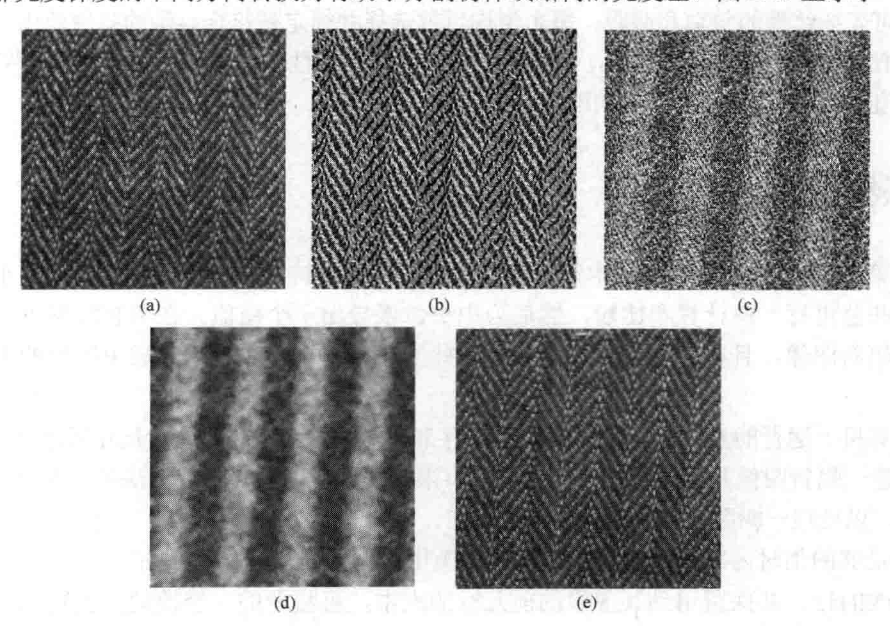


图 5.45　方向纹理处理：(a)原始的人字形图案织物图像；(b)从 0° 至 360° 标定的 Sobel 方向；(c)从 0° ～180° 并从 180° ～360° 标定的 Sobel 方向；(d)对图(c)应用中值滤波后的结果；(e)将图(d)的方向值放到图(a)的色调通道后的结果

字斜纹布上的编织式样产生垂直于图案方向的亮度梯度。当使用由梯度矢量确定的角度来指派基于从 0° 到 360° 范围的亮度值时，指向外部的梯度矢量会在每个区域中产生两个可能的结果，它们相隔 180°（或 128 个灰度值）。如果该角度改为分配 0° ～180° 和 180° ～360° 范围的灰度值，那么具有不同模式方向的每个区域会收到一个单一的、独特的亮度值范围。对该结果应用中值滤波，会消除噪声并唯一地标记这些区域。对原始图像的色调通道分配方向值表明各个区域得到了分割。

5.9　分形分析

由分形维数表征的表面粗糙度已应用到破裂面、磨损和侵蚀、腐蚀等（Mandelbrot et al.，1984；Underwood & Banerji，1986；Mecholsky & Passoja，1985；Mecholsky et al.，1986，1989；Srinivasan et al.，1991；Fahmy et al.，1991；Russ，1994）。业已证明（Pentland，1983；Peleg et al.，1984），分形表面图像的亮度模式在数学上是一个分形，它同样适用于 SEM 图像（Russ，1990a）。用来计算高程图像表面的分形维数的一种特别有效的方法是 Hurst 系数，或重标极差分析（Hurst et al.，1965；Feder，1988；Russ，1990c）。该过程将图像或表面上线性遍历的点之间的最大高度（或亮度等）差，在对数轴上画成搜索距离的函数。当最大差通过除以数据的标准差来标定时，结果线条的斜率和截距就与剖面的分形维数关联起来。

在像素级别执行这种运算很有趣，因为它允许进行局部分类，而局部分类可用于图像分割。将图像中的每个像素值转换为表示局部粗糙度（在 Hurst 系数意义上）的新亮度标定后，可采用简单的亮度阈值处理方法来分割图像。相比于测量亮度剖面中用到的一维比较方法，这种方法使用亮度变化的二维信息。

该算法检查原始图像中围绕每个位置的邻域中的像素，最大半径通常为 4～7 个像素（算上中心像素，总邻域直径为 15 个像素）。找到每种距离类中的最亮和最暗像素值，然后用它们的差来构建一幅 Hurst 图。当然，也可使用邻域中像素值的标准差。为对数（亮度变化）与对数（半径）关系执行最小二乘拟合，可表征纹理的斜率和截距，每个值均可被选择并标定到符合显示的亮度范围。与简单的邻域运算如平滑相比，这个过程很耗时，但在典型个人计算机的性能范围内仍表现良好。图 5.43(e)给出了使用截距值的结果，图 5.44(d)则给出使用斜率的一个例子。

5.10　实现说明

本章和第 4 章中讨论的许多技术都是邻域算子，它们会访问围绕每个中心像素的小区域内的像素，使用这些值进行一些计算和比较，然后为中央像素导出一个新值。在所有情形下，这个新值被用于产生一幅新图像，且在对整幅图像重复进行此操作时，下一像素的邻域中所用的值仍是各像素的原始值。

个人计算机上运行的多数图像分析系统的内存都有限，尤其是在考虑大尺寸图像时。为每个图像处理运算创建一幅新图像并不能有效地使用这种有限的资源。因此，常用方法是“在适当的位置”执行这一操作，以处理一幅图像并使用结果来替换它。

只需要足够的临时内容来容纳图像的几行。操作通常沿扫描线从左到右、从上到下来遍历图像。复制正被修改的行，并保留用到其像素的前几行的副本，可将新的（修改的）值写回原始图像内存中。行数是 $(n+1)/2$，其中 n 是邻域尺寸（如 3×3、5×5 等）。通常，从图像中复制一行所需的时间很短，它是通过许多指针来实现的，仅在移动处理到达时才需要复制每一行，然后对后续行重新使用该数组。

上面描述的一些图像处理方法使用了两个或多个中间结果。例如，Roberts 交叉算子或 Sobel 滤波器应用了其幅度随后将合并的两个方向导数。可以逐像素这样做，因此无额外的存储需求。甚至上面给出的传统操作，即在每个像素位置应用高斯平滑、拉普拉斯锐化和 Sobel 边缘检测器的操作，也只需要单次遍历图像。但在有些实现中，尤其是那些可有效编程到阵列处理器的实现中，一次会对图像中的一整行进行处理，因此为应用到每行的每个算子获得中间结果，然后为整行合并中间结果，将是最快的。保留中间结果仅需少量的额外存储空间。

实现邻域运算的另一个考虑是，如何最好地处理图像边缘附近的像素。可能的方法包括指定靠近边缘的特殊邻域规则来对一组像素排序，复制边缘处的几行像素（即假设每个边缘是一面镜子），推算边缘之外的值，或使用环绕式寻址（即假定图像的左右边缘及顶底边缘是连续的）。这些方法中没有哪一个是普遍最优的。

5.11 图像数学基础

到目前为止，本章中介绍的图像处理运算都是对一幅图像进行操作，并产生改进的结果，结果可存储在相同的图像内存中。另一类运算使用两幅图像来生成一幅新图像（可替代原始图像之一）。因为包含诸如相加、相减、相除和相乘这样的算子，因此这些运算通常称为图像算术。它们逐像素执行，故两幅图像之和包含了其亮度值是两幅原始图像中对应像素之和的像素。

还用到了其他一些算子，如比较两幅图像来保留较亮（或较暗）像素，计算绝对差，或基于邻域值的比较（例如，保留具有最大局部方差的像素，以作为合并具有不同焦距图像的手段；或按平均邻域值的比例来添加像素，以将具有不同曝光率的几幅图像混合在一起）。其他的两图像操作，如布尔“或”或“与”逻辑，通常应用于二值图像；第 8 章对此进行了介绍。

图像相加被用于对图像求平均来降低噪声，如第 4 章所示。相加运算很直接，但要明确如何处理结果。如果是两幅 8 位图像（每个像素处的亮度值从 0 到 255）相加，那么结果值的范围是从 0 到 510。若使用的是单字节整数，这就超出了图像内存的容量。一种方法是将结果除以 2，得到一幅被缩放到 0 到 255 范围的结果图像。类似的方法也应用于图像平均，其中 N 幅图像相加总生一幅总图像，然后除以 N 来缩放数据。

另一种方法是找到和图像中的最大和最小的实际值，然后动态地将结果缩放到这个最大值和最小值，以便为每个像素分配一个新值：

$$B = \text{范围} \times \frac{(\text{和} - \text{最小值})}{(\text{最大值} - \text{最小值})} \tag{5.9}$$

式中，范围是图像内存的容量，通常为 255。与第 4 章中介绍的先除以 2 再执行线性对比度扩展相比，这种方法更优，因为结果值的精度更高。整数除以 2 时，分数值会被截尾，因此可能会丢失一些信息。

另一方面，在进行动态范围确定或自动标定时，由于亮度尺度可能不一样，因此直接比较处理后的图像会更为困难。此外，自动标定耗时会较长，因为它需要两次完全遍历图像：一次用于求最大值和最小值，另一次则用于自动标定计算。出于最大化印刷物对比度目的，本书中印刷的许多图像已被自动标定。多数情形下，这一操作是作为保持精度的处理操作的一部分来执行的。

将图像加在一起会叠加信息，在某些情形下可用于创建合成图像，进而找出图像中复杂的空间关系。将原始图像与拉普拉斯算子图像或导数图像结合起来，有助于提供一些空间信息去解释来自滤波器的结果。通常，这类相加是通过改变核的中心值而直接处理的。对于拉普拉斯算子，这一修改称为锐化滤波器，如上所示。

5.12 减去图像

第 4 章中使用相减法通过删除背景调匀了图像。本章中已给出了几个相减的应用，如反锐化掩模中采用的相减，此时是从原始图像中逐像素减去平滑后的图像。初始值域为 0～255 的图像，经此类运算后，值域会变为-255～+255。通过除以 2 并加 128 来替代原始图像，或采用式（5.9）给出的用于加法的相同自动标定方法，可将数据重新标定到满足单字节内存需求。无论是固定的标定还是灵活的标定，都会有相同的优点和缺点。

有些情形下，会采用绝对差而非背景的简单相减。图 5.46 显示幻灯片上细胞的相位对比图像。正反面具有明显亮暗阴影的此类图像很难测量，原因在于两面定义边缘的规则不同，且在顶部和底部，边缘根本未显示出来，只有观看者的推测。通过使用中值滤波器消除阴影，产生了一幅无特征的“背景”图像。这幅图像和原始图像之间的绝对差显示了两类明亮的阴影。这幅图像经阈值处理后，产生了两个断开的弧，但二值图像经膨胀处理（将在第 8 章中讨论）后合并了两面。经填充和腐蚀处理后，图像变为原来的大小，清晰的细胞轮廓完全可以进行测量。

相减也是发现图像间差异的一种方法。图 5.47 显示了两幅硬币图像及它们的差。两幅图像中基本不变的部分相互抵消，并显示为均匀的中间灰度，只是由于噪声、数字化精度、光照变化等原因，有一些小的变化。差值图像中清晰地显示了两幅图像获取时所移动的硬币。暗图像中的特征不变，亮图像中的特征已消失。

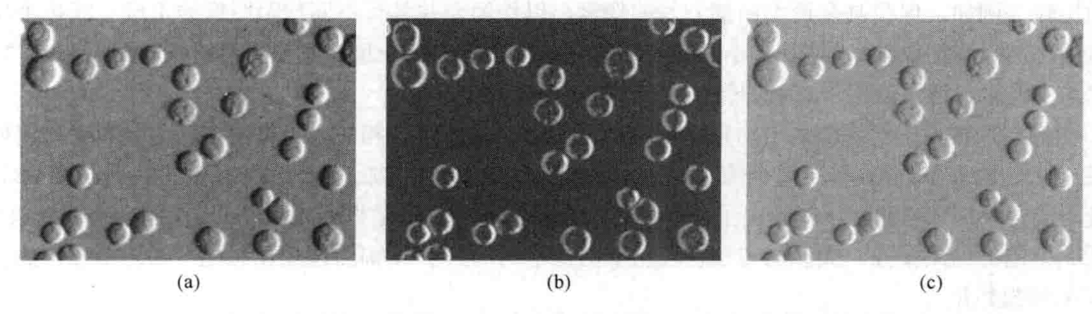


图 5.46 相位对比图像中隔离的特征：(a)原始图像；(b)原始图像和中值滤波图像间的绝对差；(c)经阈值处理、膨胀和腐蚀并叠加到原始图像上后的边界

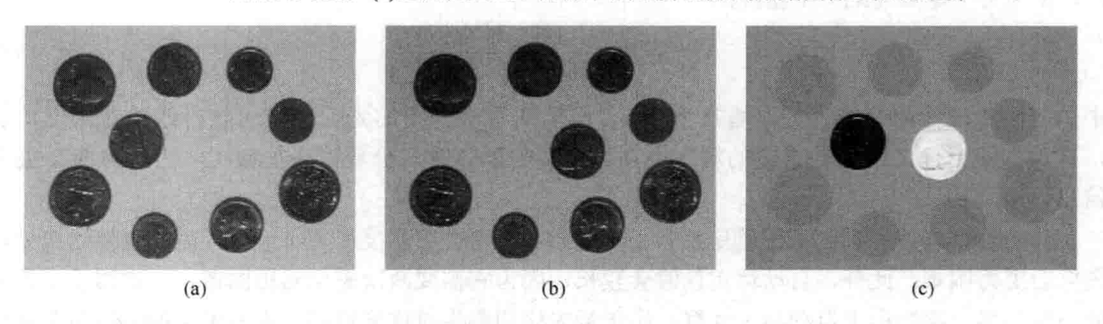


图 5.47 通过相减显示图像差异：(a)原始图像；(b)移动一个硬币后的图像；(c)逐像素相减后的差值图像

从一幅图像中减去另一幅图像，有效地从差值图像中消除了所有不变的特征，而突出那些变化的特征。如果光照和几何角度是一致的，那么像素值无变化位置的唯一差别是亮度的统计变化，这种变化是由相机或电子噪声引起的。亮暗图像中分别显示了已从视野中移除或加入到视野中的特征。

即使存在一些噪声，两幅图像相减也是我们识别不易察觉小差异的一种有效方法。图 5.48 显示了

一个例子。例中图像显示了来自 Debye-Scherer X 射线相机的两个胶片。垂直线显示了 X 射线从小样本上折射的胶片曝光，每行对应于材料结构中一个原子平面的反射。比较这些相似样本的胶片表明，多数线在位置和亮度上是类似的，因为两个标本的组成是相似的。图像中的其他微弱线则揭示了微量杂质的存在。从一组线中减去另一组线增大了相对噪声量，但显示了源自微量化合物的线，以及匹配线的亮度差别。然后可使用它们来进行测量，或用于识别目的。

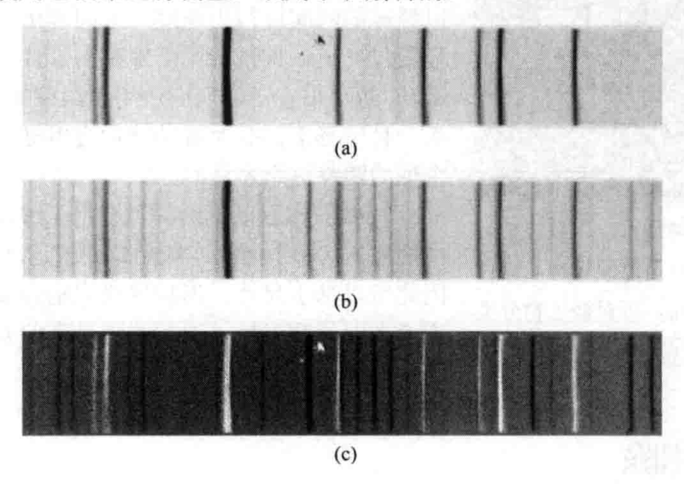
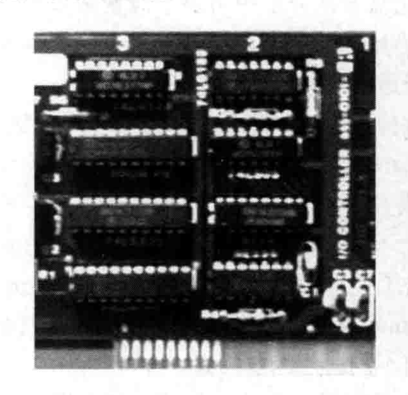


图 5.48　图像相减来增强细节的可见性：(a)和(b)源自 Debye Scherer X 射线相机的扫描图像，摄自类似的化合物；(c)图(b)和图(a)之差显示了低亮度线，它们是由样本中的微量化合物引起的

图像相减的主要用途之一是质量控制。假设获取并存储了一幅主图像，该图像显示了电路板上各元件的正确位置（见图 5.49）、包装上标签的对齐等。当从一组后续图像中减去该图像时，会强烈地突出差异，进而揭示生产中的错误。相减操作通常以专用硬件的视频帧率执行。由于希望元件精确对齐是不切实际的，因此应在相减后，根据亮和暗（失配）像素区域来指定一个容限。对侦察照片，可应用同样的技术来查看复杂场景中目的的出现或消失。在进行相减操作之前，可能需要如第 4 章那样对图像进行扭曲操作，以便对齐在不同视点拍摄的图像。



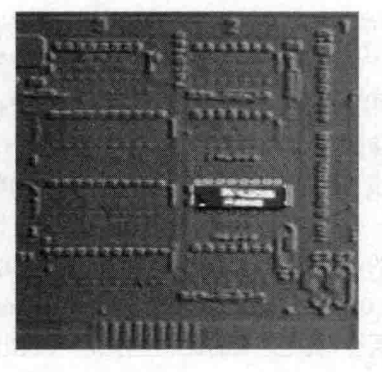


图 5.49　用于质量控制的差值图像。从每个后续元件的图像中减去主图像。在该例中，印制电路板中缺失的芯片在差值图像非常明显

天文学中使用了类似的技术。在不同时间对天空中的同一区域拍摄“闪烁”图像，是寻找移动行星或小行星的传统方式。这种技术在图像之间交替切换，以便观看者注意到两幅图像中光点的运动是明显不同的，而忽略静止的星星。已出现使用相减操作的计算机搜索方法，但对于背景噪声中的暗淡物体，还不能证明这种方法比人类观察灵敏。

在第 4 章图 4.7 所示的例子中，两幅图像的获取时间相隔 1 分多钟。两者之差清楚地显示了分针的运动。还可看到时针的很小运动，而在背对背查看这两幅图像时，是很难注意到这么小的运动的。

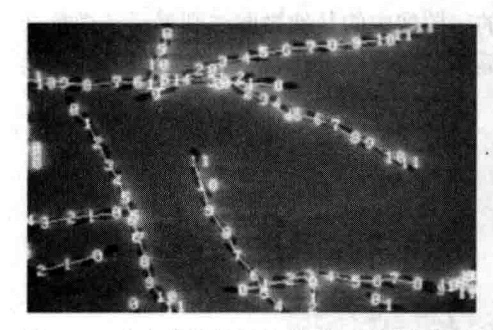


图 5.50　多个物体的运动分析。在微生物游动的交叉位置，通过假设流动路径连续且近乎直向，可对微生物分类（Gualtieri & Coltelli，1991）

如果特征足够大，且获取序列图像的速度快到足以与连续帧重叠，那么可使用相减法来测量物体的运动。此时，相减操作会给出可被测量的明亮失配区域。失配区域的长度除以经过的时间可得出速度，方向可由该区域的方向确定。在显微尺度上，该技术用于跟踪幻灯片上细胞响应化学线索的运动，也可以在动物研究中用于跟踪笼子或围栏内的运动（在后一种情形下，会对每种动物应用唯一的彩色标签，以便对个体的跟踪更为容易）。

相减操作曾以不同的尺度用于跟踪卫星照片中北大西洋中的冰川。对于两幅连续图像间的运动，这种方法非常合适，因此可以基于尺寸、形状等来识别连续图像中的相同物体。或者在路径交叉处，引起路径最小偏差的点可给出正确的匹配（见图 5.50）。

5.13　相乘和相除

图像相乘也许是用得最少的数学模式，但为了完整性，它通常会包含在提供有其他算术运算的系统中。相乘用于将一幅图像叠加到另一幅图像之上，例如，将纹理添加到表面上（通常称为“凹凸贴图”）。类似的乘性叠加，可用于向一幅反射或透射图像添加荧光或其他发射图像。

图像相乘的缺点之一是可能会产生极端值域。对于像素值域为 0～255 的 8 位图像，乘积的值域会是从 0 到超过 65000。这是一个 2 字节积，除非使用自动标定，否则只有高位字节可被存储回原来的图像内存中。因此，结果图像会出现严重的精度损失。

对于除法，数的大小也会造成问题。首先，一定要避免除以 0。避免的方法是，将所有的亮度值加 1，将值域变为 1～256 而非 0～255。此时，首先需要在分子中将每个像素乘以某个因子，来产生涵盖值域 0～255 的商，同时在值域的两端保留一些有用的精度。自动标定对于这些情形特别有用，但它不能用于需要将一幅图像与另一幅图像或校准曲线进行比较的应用中。

自动标定有用的一个相除例子是，在使用线性检测器或摄像机时移除背景（如第 4 章所示）。当需要绝对值时，相除的一个例子是计算两个或多个 Landsat 波段的亮度比（第 1 章中给出的一个例子），或者在光学显微镜下检查荧光探针时，计算两幅或多幅滤波器图像的亮度比。在荧光显微术中，时变的发射光强被归一化，方法是交替收集感兴趣行上下以不同波长通过两个或多个滤波器的图像，并校准感兴趣化合物的浓度比。在卫星影像中，亮度比（专题制图仪波段 4 为 0.5～0.6μm、波段 5 为 0.6～0.7μm、波段 6 为 0.7～0.8μm，波段 7 为 0.8～1.1μm）用于地形分类和某些岩性的识别。通过在不同的当地时间获得岩层变热或变冷的比值图像，可求出不同岩层的热惯量。

作为矿物鉴定的一个例子，硅酸盐在吸收带中显示了一个波长偏移。花岗岩、闪长岩、辉长岩和黄绿橄榄石中，硅的含量逐步降低。随着硅酸盐晶格中硅和氧原子间的键伸缩振动，吸收频带会逐渐偏移到 8～12μm 热红外波段中的更长波长处。热红外多光谱测绘仪卫星记录了此范围内的 6 个波段的图像数据，它们并合并和归一化，以找到吸收频带并识别岩层。碳酸盐岩（白云石和石灰岩）在 6～8μm 范围内有类似的吸收响应。在雷达波长，不同的表面粗糙度会在 Ka、X 和 L 波段内产生反射强度的变化，用相同的方式合并这些变化，可对表面进行测量，进而区分砂、碎石、鹅卵石和大砾石的粗糙度（Sabins，1987）。

同样，多光谱卫星影像的 0.55～0.68μm 波段（可见红光）和 0.72～1.10μm 波段（反射红外线）可用于识别植被。第一个波段记录叶绿素的吸收，而第二个波段给出树叶细胞结构的反射。比值 $(B_2 - B_1)/(B_2 + B_1)$ 消除了太阳高度差（照射角）引起的变化，并用于测量图像中的植被分布。通常，这种方法还会结合源自连续扫描的数据，进而得到光谱植被指数与时间的关系。对图像还使用了其他比值，此时的目的是测量海洋浮游植物引起的叶绿素浓度（Sabins，1987）。图 5.51 显示了这种方法的一种近似应用，它使用近红外与蓝色之比来分离出卫星影像中的植被。

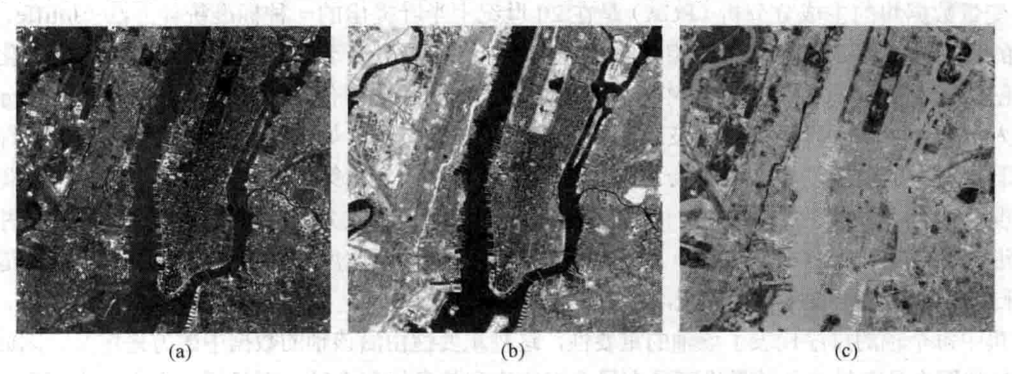


图 5.51　纽约市陆地卫星专题图像：(a)波段 1（可见蓝光）；(b)波段 4（近红外）；(c)波段 4 与波段 1 之比（显示了如中央公园这样的植被区）

比值也用于天文图像。图 5.52 显示了 NGC-2024（火焰星云，猎户座分子云团的一部分）中恒星生成区域的红外图像。红外光穿透了可阻止大部分可见光的尘埃。不同波长图像的比值或差值显示了尘埃的细节，并增强了年轻恒星的能见度。组合各个通道来进行彩色查看，是一种有用的补充。

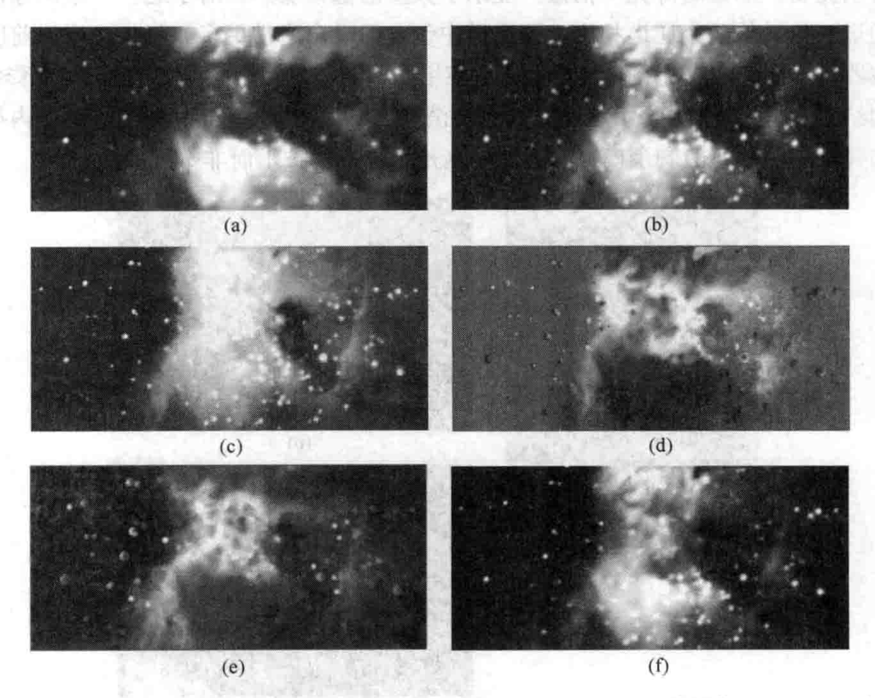


图 5.52　NGC-2024 的组合视图显示了恒星形成区域和尘埃：(a)1.2μm 红外图像；(b)1.6μm 红外图像；(c)2.2μm 红外图像；(d)2.2μm 图像减去 1.6μm 图像；(e)1.6μm 图像除以 1.2μm 图像；(f)原始的三个通道合并为 RGB 通道

5.14 主成分分析

获取多波段图像后，不论是遥感中检测到的各种可见光和红外通道，还是典型数字相机的蓝色通道，都不可能期望单个通道就可包含物体结构的最重要表示。当通道包含其他信息（如由 AFM 获取的各个信号，或可被 SEM 中 X 射线放射检测的多个元素）时，尤其如此。

多变量数据集的主成分分析（PCA）是在 20 世纪上半叶提出的一种标准统计方法（Joliffe，2002）。该方法的目的是要找到与各信号不相关的现有变量（此时指各个颜色通道）的组合，并最大化它们的方差。创建一个 $N \times N$（通道数）矩阵来保存通道间的所有协方差值。矩阵的本征向量是主成分轴。这为研究人员提供了一种将源数据轴变换为一组正交主轴的方法。如果每个像素的值被用做坐标来绘制其在色彩空间中的位置，那么主轴会由最大离散度值拟合，然后第二个轴与主轴正交，并与剩下的最大离散度值拟合，以此类推。对于一幅三通道图像，变换到主成分空间相当于颜色轴的旋转并创建三个新通道，其中第一个（最重要的）通道包含有最多的结构对比度和信息。在出现更多的通道时，原理是相同的，但可能会更难于可视化。

主集中每个轴的排序代表了该轴的重要性，这种重要性由沿该轴的数据中的方差定义。因此，第一个主轴是数据中具有最大分散量进而具有最多对比度和信息的那个轴，而最后一个主轴表示了最少量的信息。对于从一组图像中提取最大对比度和结构信息，以及找到源变量间的相关性并确定它们的贡献是降低还是冗余，是尤其有用的。除遥感应用外，该技术还未被广泛应用到成像领域（Neal & Russ，2004）。

噪声和图像人为缺陷（包括源自 JPEG 压缩的缺陷）通常留给最重要的通道。在转换回 RGB 空间前，处理最重要的通道可消除图像噪声，如图 5.53 所示。在该例中，源自消费级便携数码相机的带噪图像（在带有明亮窗户的昏暗灯光下拍摄）显示了大量的像素噪声。对于这样一幅图像，常规方法是主要对蓝色通道（与其他两个通道相比，该通道中包含的噪声通常更多）降噪，或如前所述使用一个彩色中值滤波器。变换到主成分并对最重要通道应用中值滤波器，然后再变换回 RGB 彩色空间，可得到明显更优的结果。在某些情形下，这种方法也能消除 JPEG 人为缺陷，并能消除颜色内插前芯片上所用滤色镜（如拜耳滤色片）的剩余模式。当然，这是一种专用技术而非常规应用。

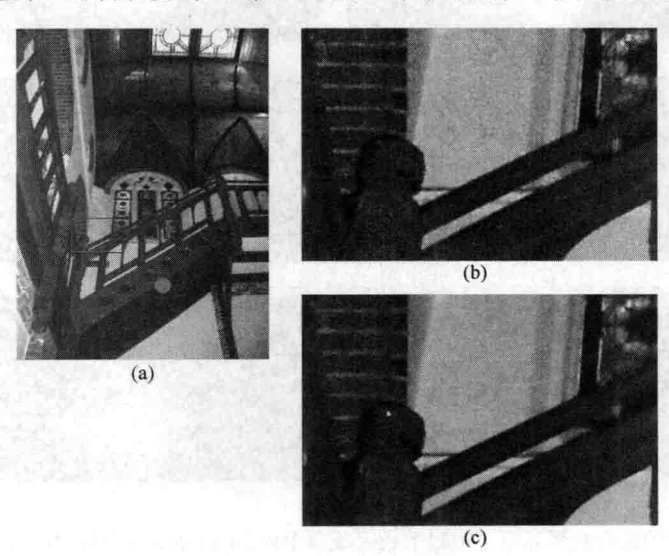


图 5.53　在 PCA 空间中降噪：(a)圈出了显示细节区域的原始图像；(b)原始带噪图像的细节；(c)经文中所述处理后的图像

检查 PCA 产生的关于源通道的信息往往是有用的。由 SEM 生成的一幅 X 射线图给出了一个例子。在这幅图像中，我们关心多个元素的分布与结合情况，因为这可给出化合物或相位信息。而且，尽管可可以挑选三个元素的组合来显示 RGB 通道，但主要目的却是在单幅图像中区分出所有的子区域。根据定义，主成分变换提供了完成这些任务的一种方法。

第 1 章中的图 1.31 和图 1.32 显示了 9 幅 SEM 的 X 射线图，每幅代表一种元素的分布与组合。此处图 5.54 再次显示了各个最初的通道。不可能以单独的颜色来显示三个以上的通道。导致这一限制的原因是，人类视觉有三类视锥，因此显示器使用了三类荧光粉（红色、绿色和蓝色）；不可能为第四种元素分配颜色（如黄色），因为黄色都是由红色和绿色组合生成的，以此类推。主成分分析可生成许多有用的图像，方法是将基本元素组合在一起，进而显示结构和相位位置。图 5.55 通过结合三个最重要的通道，显示了相位分布。方法是，将这三个通道置入彩色图像的 L、a 和 b 通道，为元素的每种组合生成唯一的颜色。

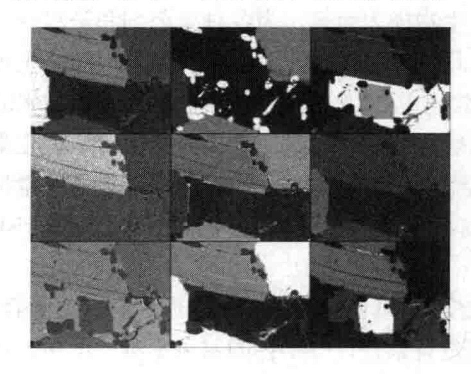


图 5.54　云母标本中 9 种不同元素的 X 射线 SEM 图（从左上角开始）：Al、Ca、Fe、K、Mg、Na、O、Si、Ti。复制自第 1 章中的图 1.32

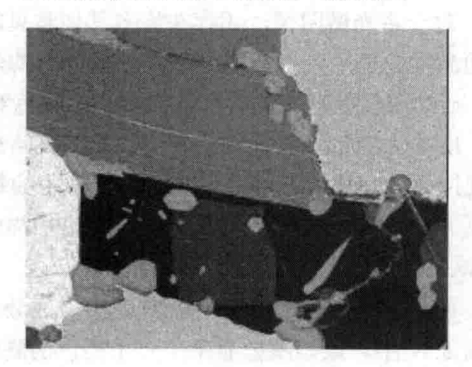


图 5.55　对图 5.54 中的 9 个通道应用主成分分析（PCA）后，组合三个最重要的通道，产生了显示出相位分布的彩色显示（与第 1 章中的图 1.33 进行比较）

主成分数据分析开始于由变换生成的协方差矩阵。表 5.4 中的矩阵给出了这些 X 射线图中 9 种元素的结果。表中的每行表示主通道（第一列）的重要性，以及每个源通道对该主要矢量的贡献。注意，用于生成图 5.55 的三个最重要通道的信息量占 9 幅原始图像中信息量的 83%以上。源通道的成分是相位组成的函数。使用共位图，可直观地显示这个数值数据。画出九维空间是困难的，但将该空间投影到任何一对轴上，则可显示出这些值的不同组合作为灰度亮度出现的频率。暗色的点群对应于原始图像中具有原始数据特征值的区域。图 5.56 所示共位图中显示的 6 个群落，表示样本中出现的相位。它们对应于图 5.55 中的 6 个彩色区域。

表 5.4　云母图像的主成分分析

		原始通道								
PCA 通道	重要性	Al	Ca	Fe	K	Mg	Na	O	Si	Ti
1	49.46%	0.2355	0.2220	−0.6026	0.0382	0.2542	0.1113	−0.0576	0.6330	−0.2228
2	20.92%	0.3970	−0.6890	0.0469	0.4005	0.1582	0.0778	0.2075	0.1687	0.3202
3	12.86%	0.0696	−0.1735	0.5571	−0.2819	0.0398	0.0488	0.2768	0.3730	−0.5963
4	7.08%	−0.3950	−0.1648	0.0958	−0.5195	0.2579	−0.0264	−0.1497	0.4030	0.5346
5	4.77%	−0.5419	0.0872	0.1360	0.5372	0.5940	−0.1062	0.0736	0.0025	−0.1489
6	2.21%	0.1996	0.6134	0.3236	0.1417	−0.0364	0.1429	0.4673	0.1924	0.4267
7	1.33%	−0.1258	0.0303	0.2784	0.4181	−0.4934	−0.0253	−0.5417	0.4388	0.0271
8	0.79%	0.4676	0.1694	0.3360	−0.0694	0.4913	0.1527	−0.5807	−0.1812	0.0144
9	0.57%	−0.2411	−0.0733	−0.0388	0.0236	−0.0572	0.9606	−0.0116	−0.0829	−0.0366

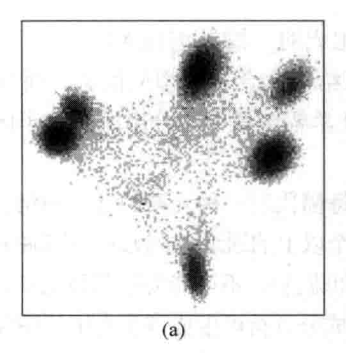
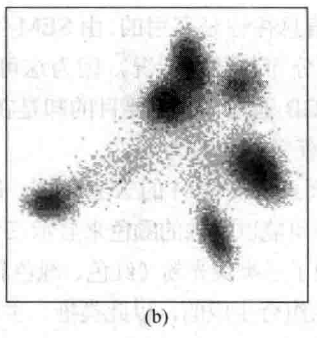
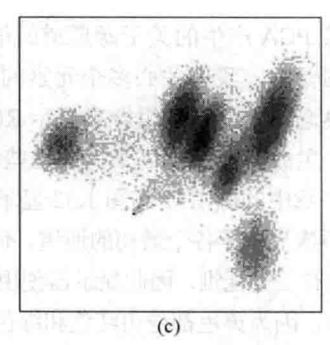

(a)

(b)

(c)

图 5.56 PCA 结果的共位图示例：(a)通道 1 与通道 2 比较；(b)通道 2 与通道 3 比较；(c)通道 2 与通道 5 比较

对于多光谱图像，还存在许多其他常见的应用。对于生物样本而言，为各种染色剂和染料使用不同的激发波长，可生成适用于主成分分析的多个源通道。此外，许多最有效的染色剂并不能产生可视化、后续阈值处理和测量所需的丰富的颜色对比度。为有效使用 PCA 技术，样品制备和成像技术的选择，应扩大到包含各种染色剂、波长、亮暗场照明及其他条件。根据定义，主成分变换会最大化最重要通道中的对比度。当成像方法为图像中的每个像素记录整个光谱（如红外、X 射线或电子能量损失信息）时，PCA 也可用于从完整的光谱扫描中提取最有意义的成分（Watanabe et al.，2009；Parish & Brewer，2010）。

即使对于常见的 RGB 图像，使用主成分来最大化对比度，也可使得对样本的处理和测量更为容易。在图 5.57 中，原始图像显示了一个染色的肠组织样本。棕色背景是许多染色的常见影响，但对该结构进行阈值处理则很困难，因为棕色中包含有红色、绿色和蓝色。对该图像应用主成分变换，增强了图像的对比度，如图所示。将生成的各个通道分配给红色、绿色和蓝色，生成了一幅细胞边界可从背景中区分的彩色图像，且细胞内部结构很明显，因此进行阈值处理非常容易。请将该结果与第 1 章中的图 1.46 进行比较，当时对该图像应用了一个任意颜色滤波器。

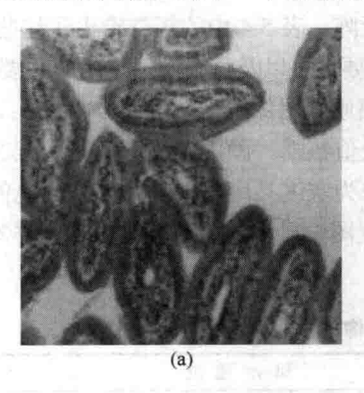
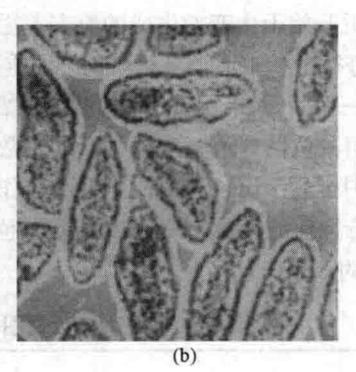

(a)

(b)

图 5.57 使用 PCA 最大化对比度：(a)原始图像（染色鼠肠光学显微照片）；(b)显示为 RGB 的主成分通道

测量染色组织最困难的问题之一，出现在染色仅产生细微差别的饱和度或亮度且彩色变化很小时。变换到另一个颜色空间（如 Lab 或 HSI）并不会产生足够的对比度或变化，进而分割出不同的结构。可视化甚至也是困难的。而使用主成分变换并将各个通道分配给显示所用的红色、绿色和蓝色通道（见图 5.58），不仅清楚地显示了结构，而且可进行测量。

图 5.59 给出了另一个应用。由于底层的印刷图案，支票背面上指纹的原始图像难以解读。应用 PCA 并将值分配给红色、绿色和蓝色显示轴，提高了能见度，而且只使用单个主要通道作为灰度图像使得指纹中的细节分析更为容易。

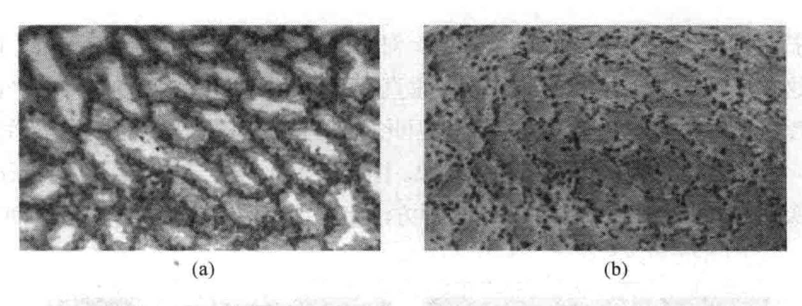


图 5.58　使用 PCA 转换细微的对比度：(a)原始图像（染色组织的光学显微照片）；(b)显示为 RGB 的主成分通道

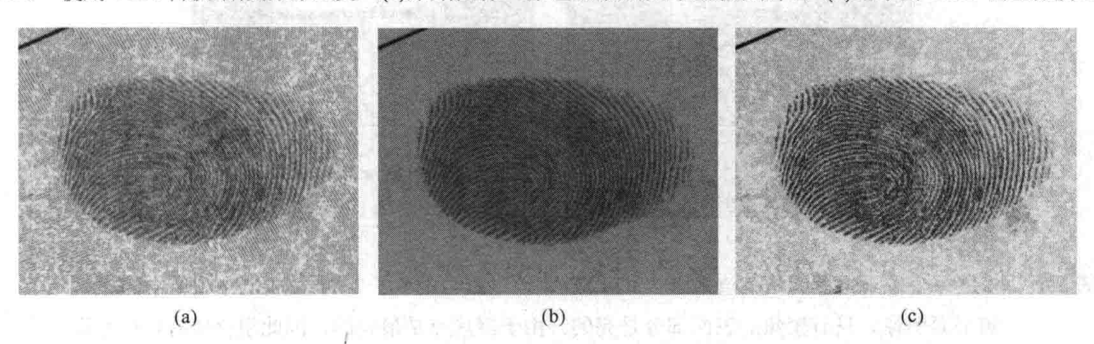


图 5.59　支票背面印刷图案上的指纹图像：(a)原始图像；(b)分配给红、绿和蓝通道的主成分；(c)主轴值作为灰度图像，具有最佳对比度

5.15　其他图像组合

图像数学还包括像素值的逻辑比较。例如，两幅（或多幅）图像可通过保留每个位置上较亮（或较暗）的对应像素来组合。偏振光显微术中存在这样一个例子，如图 5.60 所示。对样品（如矿物薄片）使用偏光镜和检偏镜，会产生这样的图像，其中有些颗粒是着色的，而其他颗粒则是深色的，它们与检偏镜的旋转有关。组合几幅不同方向的图像，并只为每个像素保留最亮的值，会生成可显示所有颗粒的一幅图像。对薄金属箔的透射电子显微图像，也可使用这种技术来组合其中有些颗粒由于电子衍射效应而变暗的图像，或消除由于箔弯曲引起的暗轮廓。采用类似的方法，也可组合粗糙物体侧光拍摄的几幅图像。

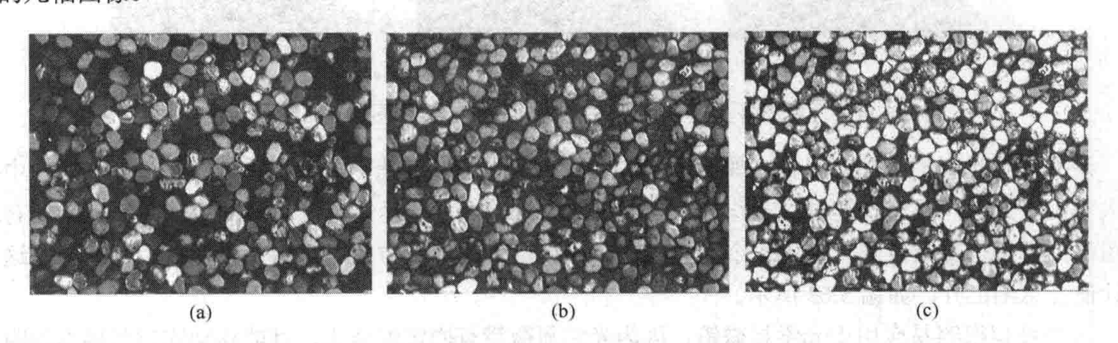


图 5.60　砂石的偏振光照片：(a)和(b)检偏镜的不同方向；(c)6 次旋转后的最大亮度图像

使用同样的方法，可创建具有大景深的共焦扫描光学显微（CSLM）图像。普通光学显微图像的景深通常都有限，原因在于透镜的孔径较大。在 CSLM 中，通过消除除焦平面中单个光照点外的任何点的散射光，检测器上的扫描光束和孔径可更多地减少这一景深。

在样本上方扫描光点，可形成一幅二维图像。对于具有不规则表面的样本而言，除焦平面上的几个表面位置外，这幅图像的其他位置都非常暗。垂直移动标本，可得到许多这样的平面图像。但在每个像素位置，光反射的最亮值出现在聚焦点上。因此，来自许多焦点深度的图像可通过只保留每个像素位置的最亮值，组合成一幅具有无限景深的图像。图 5.61 显示了一个例子。在进行比较和叠加之前，稍稍移动每幅图像，可产生另一种效果。采用这种方法组合 26 幅图像，可产生第 1 章中图 1.53 所示的表面透视图。

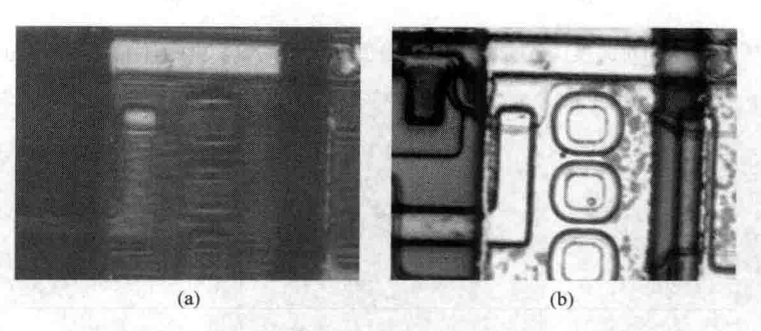


图 5.61　保留每个像素位置的最亮值来组合 CSLM 图像：图(a)显示了 25 幅集成电路图像中的一幅焦平面图像。只有聚焦的表面部分是亮的。由于聚焦点是最亮的，因此组合所有这些平面就生成了整个表面都聚焦的结果图(b)

使用传统的光学技术，也可生成大景深图像（Liu et al.，2007；Xie et al.，2007；Piper，2008）。获取一组焦平面覆盖了样本景深范围的图像后，通过在每个像素位置处保留聚焦最好图像的值，可对这些图像进行组合。图 5.62 显示了一个使用两幅图像的简单例子。在该例子中，可通过剪切和粘贴来组合两幅图像中的聚焦区域，但数学上更为可取的方式是，在这组图像的每个位置选择聚焦最好的点。

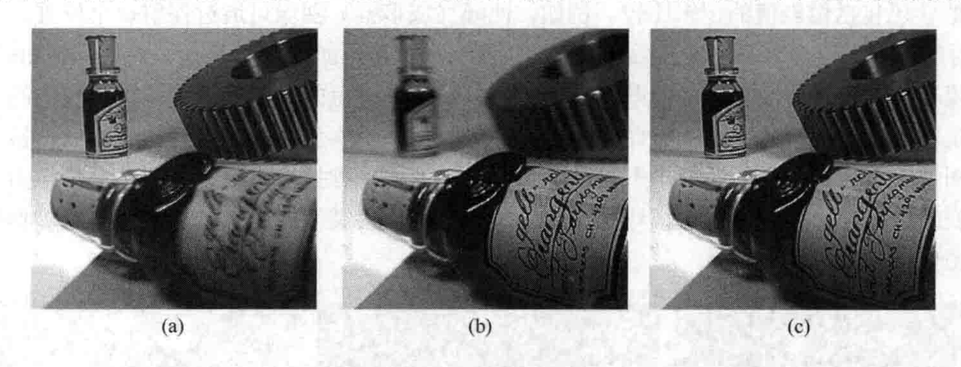


图 5.62　合并两幅图像来得到扩展的景深：(a)和(b)原始图像；(c)组合后的结果

进行这种选择的方法有多种，如第 1 章中讨论的光学自动对焦的局部自适应算法，寻找图像中的最大对比度方法，或选择高频信息作为聚焦规则的方法。一种快速且行之有效的方法是，遍历所有这些图像，在每个像素位置，使用一个小邻域来计算亮度值的局部方差。当图像聚焦时，局部方差这一统计值会急剧上升，如图 5.63 所示。

这个过程很容易应用于光学显微镜，因为光学显微镜有固定的镜头，且能移动物镜到样本的距离来对图像聚焦。它能提供一组对齐且具有相同放大率的图像。而宏观成像则很难获取一系列合适的图像。而受限于几何结构的相机即使严格控制，也不能通过调整光学镜头来得到正确聚焦的图像。相反，使用长焦镜头，可通过改变与物体的远近，来得到正确聚焦的各幅图像。但在这些条件得到满足后，就可得到无限的景深，如图 5.64 所示。

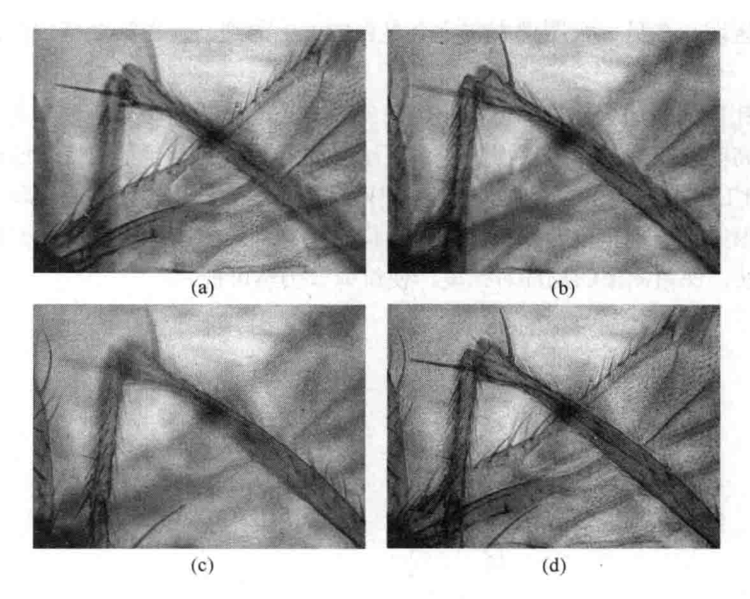


图 5.63　使用显微图像扩焦：(a)、(b)和(c)显示苍蝇眼睛和腿的焦平面图像；(d)使用局部方差最大法从每幅平面图像中选取像素值的合并结果

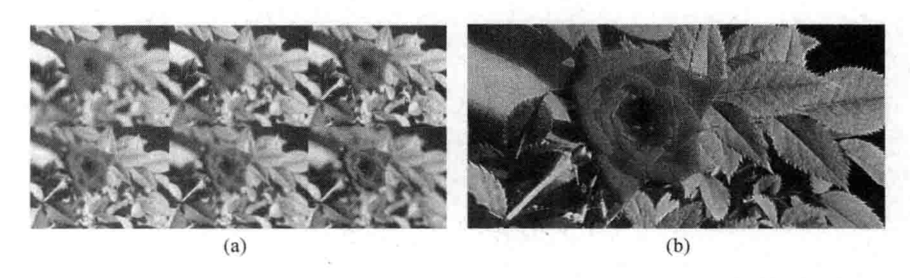


图 5.64　合并一系列宏观图像来得到扩展的景深：(a)改变相机镜头到物体的距离所得到的图像系列；(b)合并后的结果

在图像中出现极端的亮度级时，也需要合并同一场景或物体的多幅图像。取决于相机传感器的阱尺寸和数字化的位深，在单幅图像中可能不能捕获场景的全部范围亮度。胶片相机也受限于可被记录亮度的最大范围。如果场景是静止的，那么使用不同的曝光时间获取两幅或多幅数字图像可提供一种解决方法。在图 5.65 所示例子中，这样的两幅图像可通过检查围绕每个像素位置的一个小邻域的平均亮度并计算各个场景的加权平均来组合，以便替代被剪切为黑色或白色的值，维护局部对比度，进而保留整个场景的外观。这就使得我们可以看到两幅图像中阴影区域（在一幅图像中它完

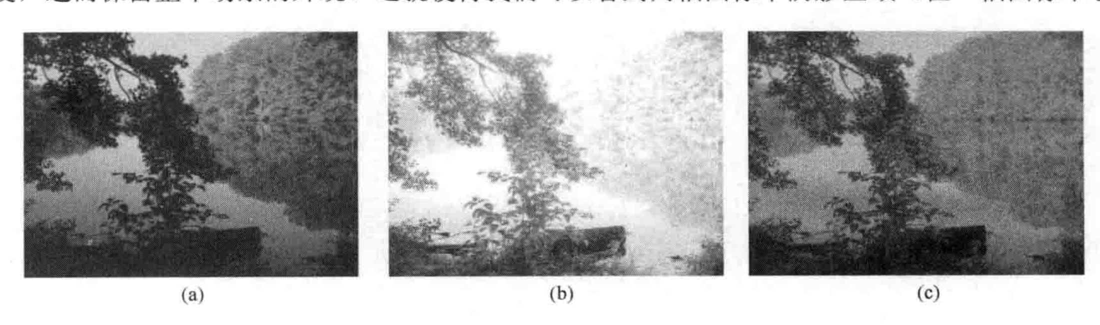


图 5.65　合并不同曝光时间的两幅图像：(a)和(b)原始图像；(c)合并后的结果

全是黑的）和明亮区域（在另一幅图像中它完全是白的）中的细节。在第 2 章中的图 2.8 显示了另一个例子。

具有大范围对比度的图像会出现另一个不同的问题。无论这些图像是通过合并多幅图像获得的，还是通过使用具有高动态范围的相机获得的，由于输出设备的限制，显示或打印这样的图像都会很困难。与用于 CRT 和 LCD 显示相比，这一限制对硬拷贝打印更为严重，但与能解释场景的人类观察者相比，两者都不能很好地处理亮度范围。因此，人们开始寻求压缩图像整个亮度范围的方法，即使用上面给出的几种方法，保留图像的局部对比度，使得细节仍然可视。

第 6 章　频率空间中的图像处理

6.1　关于频率空间

关于图像分析的书籍都会花很大的篇幅来介绍傅里叶变换（尤其是 Pratt，1991；Gonzalez & Wintz，1987；Jain，1989）。原因之一是，在频率空间进行某些图像处理及测量运算的效果较好。许多相同的操作也可对空间（像素）域图像进行，但需要更大的计算代价。原因之二是，这些图书的作者通常具有电气工程和信号处理的背景，他们对其他类型信号的数学处理方法更为熟悉，尤其是现代电子学中的一维（时变）电信号。

但热衷于为增强或测量目的而对图像应用计算机方法的典型图像分析人员，面对成页的数学运算（和符号表示）时并不轻松。此外，他们也不能很好地将这些概念与典型图像分析软件的操作关联起来。由于不了解所讨论主题与现实生活中所遇图像问题间的联系，潜在用户更可能会忽略这一主题。无疑，这样做并不可取，因为使用频率空间方法能为许多实际应用带来益处，并能提供另一种方式来解释和理解图像的内容。获取这些技术的实用知识，并不需要我们深入地了解数学运算。

由于种种原因，傅里叶变换和其他频率域空间变换被应用于二维图像。有些变换与增强可视性及选择感兴趣特征或结构进而进行测量的目的关系不大。例如，有些变换方法用于图像压缩，压缩后图像与原始图像相比，所需的存储空间更少，因而更便于传输。在这类应用中，为观看目的需要重建图像（将其从频率域转换回空间域）。因此，需要在尽可能不降低图像质量的情况下，快速完成正向和反向变换。

图像质量是一个难以捉摸的概念，但肯定包括亮度级和色彩值的改变、特征边界的定义和位置、图像中细小纹理和细节的引入与消除。一般来说，压缩程度愈高，图像失真愈严重，如第 3 章中所述。多数肉眼使用的压缩和还原图像所能容许的图像退化量，要远大于图像分析目的所能接受的图像退化量。

本书所涉的多数情形并不关心图像的传输，因此本章忽略图像的压缩，并假定变换保留了所有数据。由于变换会对图像信息完整地编码，因此会以所用计算机的算术精度精确地重建图像（这种算术精度通常要高于原始图像传感器或模数转换器的精度）。

虽然可用的图像变换类型很多，但最知名的是傅里叶变换。部分原因在于计算傅里叶变换存在强大且有效的算法，即后面将要介绍的快速傅里叶变换或 FFT（Cooley & Tukey，1965；Bracewell，1989）。虽然有些程序计算时实际上采用的是等效快速哈特利变换或 FHT（Hartley，1942；Bracewell，1984，1986；Reeves，1990），但计算所得的频率域空间表示通常与傅里叶方法产生的相同。为解释方便，本章将介绍最为知名的傅里叶变换。

推导傅里叶变换数学原理的常用方法是，从一维波形开始，然后扩展到二维（图像）情形。原理上，这种方法可扩展到三维或更多维的情形，尽管三维或更多维的情形更加难以显示。例如，在某些层析图像重建中，会用到空间域（此时立体图像由体素而非像素构建）和三维频率空间表示间的三维变换。

本章会尽可能简洁地说明傅里叶变换的数学推导过程，有“积分恐惧症”的读者可略过这部分内容而直接阅读后面的例子，仅在需要深入理解时再回头阅读这部分内容。

6.2 傅里叶变换

通过使用相当标准的术语和符号，我们以函数 $f(x)$ 开始，其中 x 是一个实变量，表示时间或图像中某个方向上的距离。我们通常称该函数为空间域函数或时间域函数，而称下面介绍的变换 F 为频率空间函数。函数 f 是一个连续且行为良好的函数。在一幅数字图像中，x 的值不是连续的，而是离散的（基于像素间距），亮度值也是量化的，所以这些值是存在于计算机之外的真实或模拟图像的采样。

傅里叶定理表明，任何良态一维函数 $f(x)$ 都可构建为一系列递增频率的正弦项和余弦项。函数 $f(x)$ 的傅里叶变换记为 $F(u)$，并指定了必须加起来形成 $f(x)$ 的频率项 u 的个数。它可写为

$$F(u)=\int_{-\infty}^{+\infty} f(x)\cdot \mathrm{e}^{-2\pi \mathrm{j}ux}\mathrm{d}x \tag{6.1}$$

式中，j 照常是 $\sqrt{-1}$。指数记号的使用依赖于数学恒等式（欧拉公式）：

$$\mathrm{e}^{-2\pi \mathrm{j}ux}=\cos(2\pi ux)-\mathrm{j}\sin(2\pi ux) \tag{6.2}$$

该变换的一个非常重要的特性是，给定 $F(u)$，就可以同样的方式（注意指数中符号的变化）恢复空间域函数 $f(x)$：

$$f(x)=\int_{-\infty}^{+\infty} F(u)\cdot \mathrm{e}^{2\pi \mathrm{j}ux}\mathrm{d}u \tag{6.3}$$

式（6.1）与式（6.3）一起构成了正向和反向傅里叶变换。函数 $f(x)$ 是一个实函数，如时变电压或空间变化的图像亮度。但变换函数 $F(u)$ 通常是复数，即它是实部 R 和虚部 I 之和：

$$F(u)=R(u)+\mathrm{j}I(u) \tag{6.4}$$

上式用极坐标而非笛卡儿坐标可更方便地写为

$$F(u)=|F(u)|\cdot \mathrm{e}^{\mathrm{j}\varphi(u)} \tag{6.5}$$

式中，$|F|$ 表示振幅，φ 表示相位。振幅的平方 $|F(u)|^2$ 通常被称为功率谱，或 $f(x)$ 的谱密度。

从负无穷到正无穷的积分实际上可以简化为递增频率项的求和，项数受限于图像中采样点的有效间距。离散傅里叶变换可写为

$$F(u)=\frac{1}{N}\sum_{x=0}^{N-1} f(x)\cdot \mathrm{e}^{-\mathrm{j}2\pi ux/N} \tag{6.6}$$

式中，N 取决于函数 $f(x)$ 的采样点数，采样被认为是等间距的。同样，逆变换与此类似（但不完全相同，它缺少 $1/N$ 项且指数的符号有变化）：

$$f(x)=\frac{1}{N}\sum_{u=0}^{N-1} F(u)\cdot \mathrm{e}^{\mathrm{j}2\pi ux/N} \tag{6.7}$$

从 0 到 $N-1$ 的频率值 u 表示加在一起构成函数 $f(x)$ 的离散频率分量。如在连续情形下那样，$F(u)$ 是复数，可写为实部和虚部形式，也可写为振幅与相位形式。

求和通常会对多达图像一半尺寸的各项执行（以像素为单位），因为它至少需要两个像素亮度值来定义出现的最高频率。这一极限就是奈奎斯特频率。因为求和所含项数是原始图像宽度的一半，但每项都有一个实部和一个虚部，那么由傅里叶变换产生的数值总数，就与原始图像宽度的像素数相同，或与时变函数的样本数相同，所以不存在压缩。由于原像素值通常是小整数（例如，对 8 位灰度图像为 1 字节），而傅里叶变换所产生的值是浮点数（在最佳实现中是双精度数），因此数据需要更大的存储空间。

在连续和离散两种情形下，将一维函数直接扩展到二维（或三维）函数的方法如下：用 $f(x,y)$ 或 $f(x,y,z)$ 代替 $f(x)$，用 $F(u, v)$ 或 $F(u, v, w)$ 代替 $F(u)$，并对两个（或三个）变量而非一个变量求和或积

分。由于 x, y, z 是正交的，因此 u, v, w 也是正交的。这意味着变换可以在每个方向上单独进行。例如，对于一幅二维图像，可先对图像的每一行进行一维变换，为每个像素生成一个具有复杂数值的中间结果，然后对每一列进行一维变换，最终生成期望的二维变换。

下面的程序说明了计算函数 FFT 的方法。它是用 C 语言编写的，但可转换成其他任何语言。对子程序进行输入后，数组数据就容纳了将被变换的值，数组按实部和虚部进行排列（这些复数的虚部通常初始为 0），且数组大小是 2 的整数次幂。变换以同一个数组返回。第一个循环对输入数据进行重排，而第二个循环则逐次加倍，这是 FFT 方法的核心。对于逆变换，必须改变角度 θ 的符号，且最终值必须按点数缩放。

```
// data[]中的偶数位置是实部，奇数位置是虚部。
// 输入必须是 2 的整数次幂
void   FFT_Cooley_Tukey(float* data, unsigned long size) {
        unsigned long mmax, n = size << 1;
        unsigned long i, j = 1;

    for (i = 1; i < n; i += 2) { //按时间抽取
        unsigned long m;
        if (j > i) { //复数交换
                swap(data[j-1], data[i-1]);
                swap(data[j], data[i]);
        }
        m = size;
        while ((m > = 2) && (j > m)) {
                j -= m;
                m >>= 1;
        }
        j += m;
}
mmax = 2; //Danielson-Lanczos, 分治
while (mmax < n) {
        unsigned long istep = mmax << 1;
        float           theta = -(2*3.141592653589793/mmax);
        float           wtemp = sin(0.5 * theta);
                        //复数: W = exp(-2*pi*m/levelp)
        float           wpr = -2.0 * wtemp * wtemp;
                        //W 的实部
        float           wpi = sin(theta);
        float           wr = 1.0;
        float           wi = 0.0;
        for (m = 1; m < mmax; m += 2) {
                for (i = m; i <= n; i += istep) {
                        unsigned long k = i + mmax; //第二个序号
                        float tempr = wr * data[k-1] - wi * data[k];
                        float tempi = wr * data[k] + wi * data[k-1];

                        data[k-1] = data[i-1] - tempr;
```

```
                    data[k] = data[i] - tempi;
                    data[i-1] += tempr;
                    data[i] += tempi;
                }
                wtemp = wr;
                wr += wr * wpr - wi * wpi;
                wi += wi * wpr + wtemp * wpi;
            }
            mmax = istep; //到下一级
        } //mmax 同时继续增加
    }//FFT
```

将这个一维变换应用到二维图像的每一行和每一列，并不是执行这一计算的最快方法，但是到目前为止最容易理解并在许多程序中用到的最简方法。稍快的一种方法称为蝶形运算，因为这种利用整个二维图像的各组像素对值，产生了相同的结果。将一组 W 值存储为预先确定的常数，也能提升计算速度。许多软件数学库函数中包含有高度优化的 FFT 程序，有些程序甚至允许数组大小不是 2 的整数次幂。

原始图像到频率空间的变换结果，在每个像素处是复数值。以易于解释的方式，很难显示变换结果。多数情形下，只能显示出振幅而忽略相位。如果使用振幅的平方，则称为图像的功率谱，因为不同的频率表示为到原点的不同距离，不同的方向则表示原始图像中的不同取向，每个位置的功率显示了图像中该频率和方向出现的多少。这种显示对于检测和隔离周期性结构或噪声特别有用，这将在下面介绍。但功率谱本身不能用来恢复原始图像。还需要相位信息，尽管很少显示相位信息，且难以直观地对其进行解释。振幅或功率值通常会涵盖非常大的数值范围，因此许多系统改为显示值的对数，而这也是此处所用的约定。

6.3 简单函数的傅里叶变换

在关于傅里叶变换的数学教材中（通常处理一维情形），常见的说明方法是，使用傅里叶级数展开中的有限项之和来拟合一个任意函数的质量。组成各级数正弦项的参数是振幅、频率和相位，如图 6.1 所示。图 6.2 显示了台阶函数的情况，给出了由一组正弦波叠加后产生理想台阶的能力。台阶函数在图像中很重要，因为一个亮度台阶通常对应于一个物体或结构的边缘。傅里叶变换的重要特征之一是，级数展开中的各项是独立的，项数增加，拟后的质量也会提高。

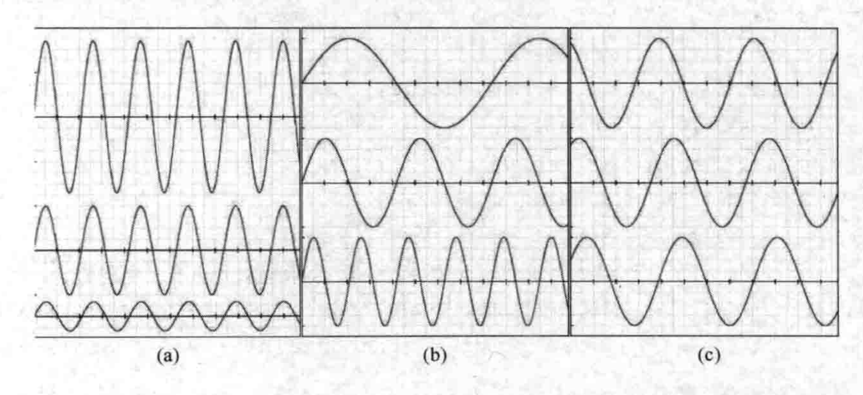


图 6.1　正弦波 $y=$ 振幅 $\times\sin$（频率$\times 2\pi x+$ 相位）的参数：(a)振幅；(b)频率；(c)相位

假设该函数是重复的或周期性的。拟合通过该区间的右端和左端，就好像函数在两个方向上无止境地重复一样。二维情形下同样如此；空间域中的图像实际上是一小块无限重复的图案。如果图像的左右边缘或上下边缘不同，那么会对变换结果产生显著影响。一种解决办法是，把该图像嵌入到一个由零值或像素平均亮度值组成的更大图像中。这种“填充”也允许将 FFT 程序应用到大小不正好是 2 的整数次幂（64, 128, 256, 512, 1024, …）的图像上，但会使图像大小在每个方向扩展两倍，且要求 4 倍的存储和计算量。如下所示，当对不同大小的图像进行相关时，填充特别有用。

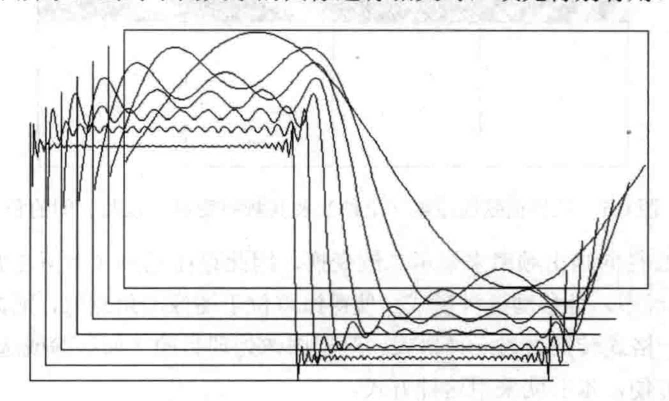


图 6.2　在傅里叶级数中用 1、2、3、5、10、20 及 40 项对台阶函数的逼近

图 6.3 画出了图 6.2 所示傅里叶变换的傅里叶系数的振幅与频率的关系。对于台阶函数，它由奇数项组成。作为傅里叶变换的典型情况，前几项（低频）的振幅较大，而高阶项（高频）的振幅较小。图 6.4 重申了傅里叶变换处理的对偶性。空间域和频率域以完全不同的方式显示了信息，但这些信息是相同的。当然，频率变换的振幅图并未显示重要的相位信息，这些值实际上被理解为复数。移动空间域图像不改变振幅值，而确实会改变每个正弦分量的相位值。

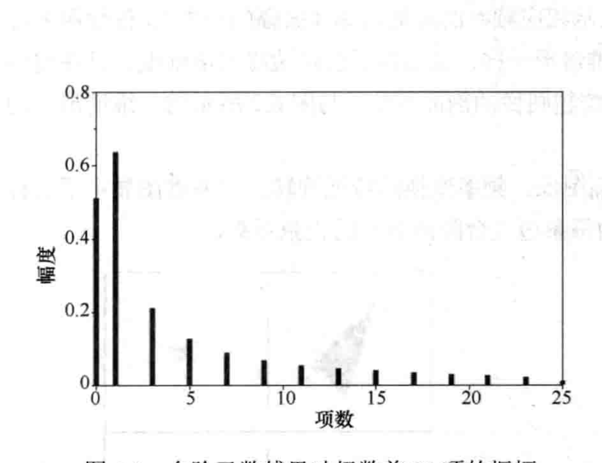


图 6.3　台阶函数傅里叶级数前 25 项的振幅

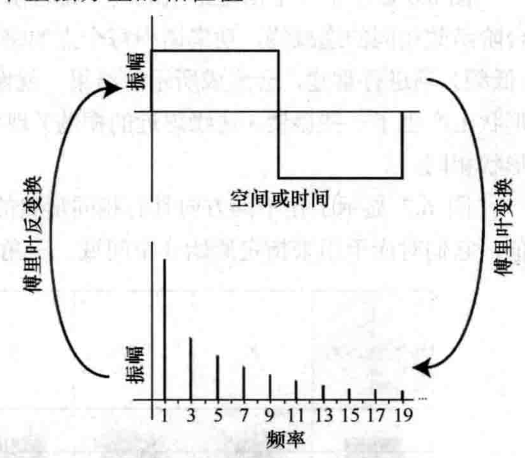


图 6.4　台阶函数正反变换的作用及其空间域和频率域的表示

在研究这些变换时，再次明确坐标轴表示频率是非常重要的。低频项决定了函数的整体形状，而高频项则用来锐化边缘并提供细节。这些项必须彼此无关［等效于声明基函数（正弦波）彼此无关］。将变换求和扩展到包含高频系数，并不会改变这一事实，而选择任何特定范围内的项来重建该函数，则使用这些频率可得到最大的精度。

关于二维情形，图 6.5 显示了亮度正弦变化的 4 幅图像。前三幅图像中存在间距（频率）和方向变化，第 4 幅图像则是前三幅图像之和。对每幅图像来说，二维频率变换特别简单。每个“纯音”都

有由单点（识别频率和方向）组成的一个变换。由于成图坐标约定，该点显示在围绕原点的两个对称位置，原点则位于功率频谱图的中心位置。和图像的变换刚好是各幅图像变换之和。

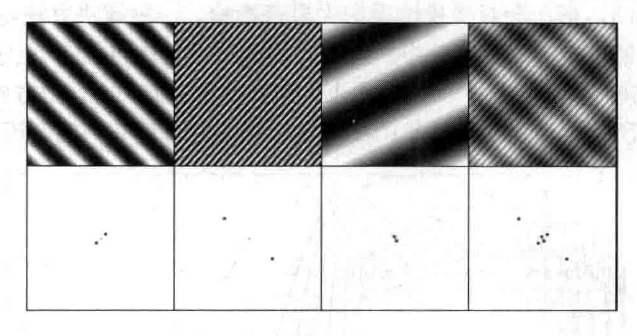


图 6.5　三种正弦线图案（上部）及其频率变换，以及它们的和

通常以原点为中心径向画出频率来显示二维变换，因此存在冗余（上下左右是重复的，关于原点旋转对称）。在某些实现中，会移动这种模式，使得原点位于图像的角点处，而高频位于中心。通过交换显示象限，可将一种格式转换为另一种格式。对于图像处理目的（如移除或选择特定频率等），使原点位于中心位置最为方便，本书就采用这种方式。

使用极坐标最容易说明二维功率谱。频率随半径 ρ 的增加而增加，方向取决于角度 θ。低频（最大间隔）正弦的功率谱有一个靠近原点的点，高频正弦具有更远的点，并在识别这些线方位的方向上。原点的值称为直流电平，表示原始图像的平均亮度。图 6.5 中三个正弦波之和生成了一幅图像，该图像的频率变换是各幅图像变换之和。对于下面的许多讨论而言，这种相加原则很重要。从频率变换中的某个位置删除信息，相当于从空间域图像的每部分删除相应的信息。

图 6.6 显示了一个由矩形组成的二维台阶。该图像的二维频率变换在 x 轴和 y 轴方向形成了与一维台阶函数相同的递减峰。功率谱中每个点的暗度表示相应频率的对数功率（振幅的平方）。仅使用中心（低频）项进行重建，会形成所示的结果。就像一维情形一样，这限制了台阶边缘的清晰度，且在整体形状上产生了一些振铃（边缘附近的振荡）现象。跨越图像的剖面形状，与图 6.2 所示的一维情形下的形状相同。

图 6.7 显示了在不同方向具有相同形状的两幅图像。频率变换随特征旋转。功率谱图显示了三行值，它们对应于用来指定原始（空间域）三角形的每条边（台阶函数）的正弦系列。

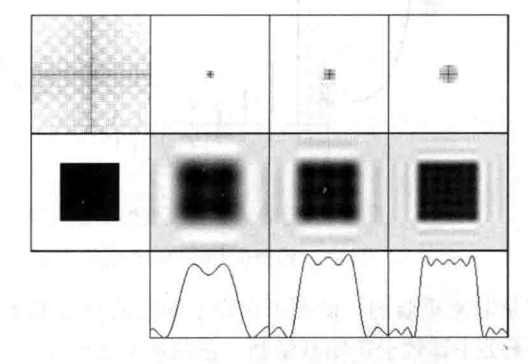


图 6.6　左侧是一个二维台阶函数及其频率变换（顶部）。每个窗格显示以不同项数的重建（显示为频率变换的一部分）。底部一行显示了过重建空间域图像中心的水平线剖面

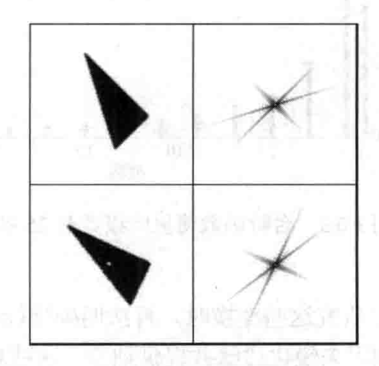


图 6.7　空间域图像的旋转（左）及频率变换的对应旋转（右）

6.4 频率和方向

使用简单图像导出我们熟悉的频率空间变换的功率谱显示，是有帮助的。在图 6.8(a)中，这些线可用一个单峰来代表，因为它们的亮度曲线是完美的正弦形式；因此只出现单一频率。如果线剖面不同，如更像方波而非正弦波，则表示该形状需要更多的项，进而在功率谱上会出现频率递增的多个峰，这称为谐波。图 6.8(b)显示了一个例子，其中频率变换是由多个最低频率处的一系列峰构成的，这些峰具有相同的走向（垂直于线的走向）。

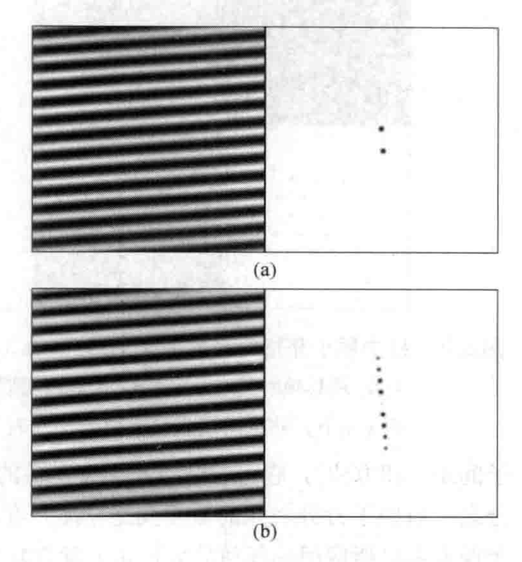


图 6.8　一组线（左）及其频率变换（右）：(a)正弦线；(b)具有非正弦亮度剖面的线

图 6.9 显示了使用一幅真实图像的例子，它是染色骨骼肌的光学显微图像，由于参差不齐的染色对比度（甚至对比度扩展），存在刚好可见但难以测量的带间距。傅里叶变换给出了可识别这种带状结构的点。通过测量到第一个主峰的距离，可以很容易且准确地求出间距。仅使用第一个峰的重建给出了基本的带间距。添加二次和三次谐波后，可更好地定义条带形状。这就引入了利用掩模或滤波器来选择（或消除）某些频率以重建一幅图像的思想，详见下文。

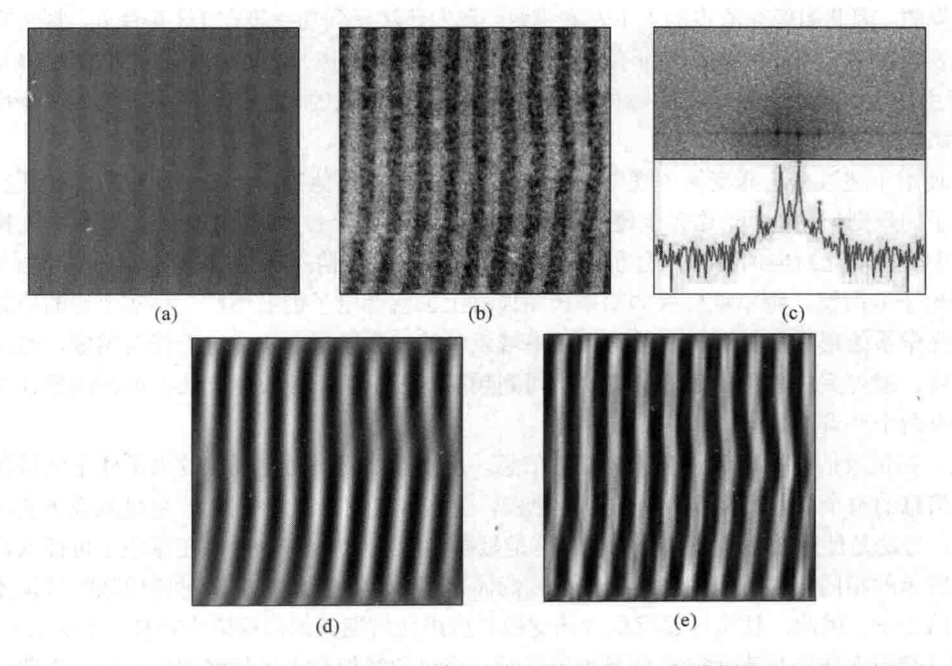


图 6.9　染色（甲苯胺蓝）骨骼肌的 1μm 截面光学显微图像：(a)原始图像；(b)对比度扩展后的结果（由于染色对比度参差不齐，条带间距仍难以测量）；(c)FT 功率谱，显示了可识别条带间距的点，以及振幅与频率关系曲线图（标出了谐波峰）；(d)只利用第一个谐波的重建结果给出了基本的条带间距；(e)添加二次和三次谐波后，改善了条带形状

上面的例子表明，原始空间域图像中的任何周期结构，可由功率谱图像中的一个峰来表示，该峰的半径对应于频率或间距，方向对应于走向。在由多数非周期信息组成的一幅真实图像中，任何这样的峰都叠加到一个宽背景上，有时叠加到噪声背景上。然而，与寻找原始的周期结构相比，寻找这样的峰通常更容易一些，且与试图从原始图像中提取相同的信息相比，精确地测量峰的位置要更容易且更准确，因为在频率域中所有这些信息均被有效地平均。

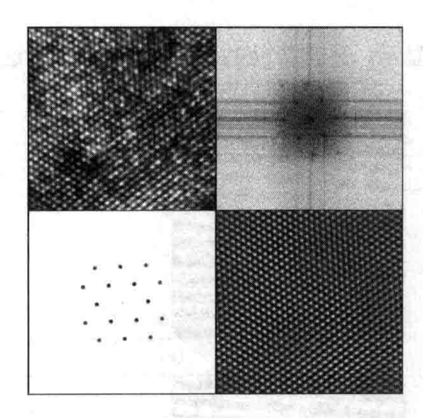


图 6.10　硅中原子晶格的高分辨率 TEM 图像（左上），及其频率变换（右上）。仅隔离周期频率（左下）并重变换只显示出原子（右下）

图 6.10 显示了峰位置测量的一个二维例子。这幅空间域图像是纯硅中晶格结构的高分辨率 TEM 图像。规则的亮点间距表示晶格的原子结构。测量这些原子间的各个间距非常耗时，且不是特别准确。该图像的频率域表示清楚地显示了周期性。这一系列峰表明，亮度变化不是简单的正弦波，而是包含了很多更高次的谐波。一次峰给出了基本的原子间距（和方位），它可以通过将功率谱中的峰位置内插到像素宽度的几分之一来测量，像素宽度的几分之一对应于万分之几的原子间距精度。在功率谱中产生该点的特征间距，是图像宽度（例中为 256 个像素乘以所应用的任何放大校正）除以功率谱中原点到峰的像素距离。

该图中还显示了可在许多 FFT 功率谱图像中找到的几个其他特征。暗色垂直线和水平线对应于所需的序列项，因为图像的左右边缘和上下边缘不匹配。由于傅里叶变换的数学运算假定被变换函数是连续且无界的，因此图像在原点的上下左右重复。我们还注意到功率谱在 1/3 半径处，振幅有相当大的突降。这表明在水平方向的高频处没有信息（只有随机像素噪声），并揭示了原始图像的实际分辨率。如第 1 章所述，许多图像是用比实际分辨率更多的像素来记录的，且如下所示，傅里叶变换可用于测量该分辨率。

如果使用下述过程选择变换中规则间隔的峰而抑制所有其他频率，那么图中所示逆变换只有原始图像中原子的规则周期间距。这完全遵循了图 6.5 中给出的可分性与可加性原则。从频率变换中去掉与原始图像中随机变化对应的各项，会在反变换图像中删除这些信息，并只显示那些周期性结构。

对于电子显微镜，频率域变换的功率谱图像看上去就像电子衍射图样，事实上也的确如此。使用显微镜的光学系统形成衍射图样，是计算频率域表示的一种模拟方法。对于任何图像，通过建立合适的光学系统，就可完成这种操作。虽然这是得到频率域表示的一种快速方法，但在图像处理中使用这种方法会有两个严重的缺陷。

首先，在记录衍射图样时，会丢失相位信息，所以不能由衍射图样照片来重建空间域图像（记录包括相位信息的整个衍射图样会得到一幅全息图，由它可能重建原始图像）。由显微镜下的衍射图样可执行重建，方法是使用合适的镜头（事实上这是显微镜的工作原理），所以在原理上可插入许多掩模来产生与图所示的相同结果。但生成这些掩模或滤波器很困难，且通常必须对待增强的每幅图像单独地进行严格的工作。因此，使用计算机执行该变换并应用任何期望的掩模更为容易（且更为可控）。

使用计算机对频率域表示进行测量也很容易。通过曲线拟合来定位峰的中心时，需要记录衍射图样（这会在许多图样的较宽动态范围上引入非线性或饱和）。考虑记录空间域图像、计算频率变换、交互式或自动测量的速度，通常也会选择计算机来进行这种处理。这一分析很容易实现，方法是调节显示对比度，以便可在功率谱中看到亮点和暗点，并对功率谱应用像背景调匀这样的图像处理工具和顶帽滤波器来隔离出感兴趣的峰。

当来自周期性结构上的点叠加到普通的背景上时，作为整个频率变换中总功率一部分的这些点的总功率，可定量测量结构中的周期性程度。通过比较功率谱中各值之和，它也可用于比较不同的周期性（不同的间距或方位）。对于电子衍射图样，它是各个平面的原子密度与原子散射剖面的函数。

虽然显示的功率谱对应于一个衍射图样，且是频率空间信息的最熟悉展示，但不要忘记还需要相位信息来重建原始图像。图 6.11(a)显示了一幅由规则的点图案组成的测试图像。图 6.11(b)和图 6.11(c)分别显示了振幅和相位值（相位信息往往不显示，因为它在视觉上很难解释）。如果在傅里叶变换中删除了相位信息，通过将所有相位值设为零，重建结果［见图 6.11(d)］显示出了一些相同的周期性，但物体不再可被识别。定义图像所添加的各个正弦波已发生相移，因此单个特征边界未被重建出来。

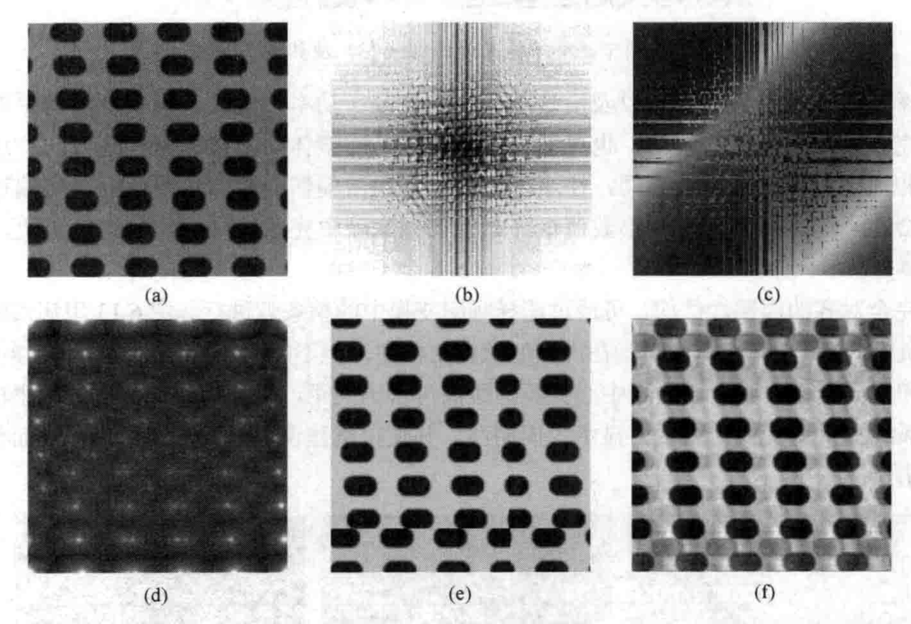


图 6.11　说明相位作用的测试图像：(a)原始规则图样；(b)图(a)的傅里叶变换功率谱；(c)变换中的相位值，它是通过使用彩色来表示从 0° 到 360° 的角度来显示的；(d)通过将相位信息设置为零后，图(b)的反变换；(e)带有任意空间移位以在原始图像边界处显示不连续性的原始测试图像；(f)图(b)的重新变换结果，水平和垂直线的振幅已减小到零，因此图像的左右边缘及上下边缘已强制匹配

假设图像是一个无限序列的一次重复也很重要。多数真实图像的左右边缘或上下边缘并不完全匹配。因此，在边缘处会产生一个大的台阶函数，如果图像出现任何偏移，台阶会更加明显［见图 6.11(e)］。尽管相位值被移动，但这不会改变功率谱图像。这种不连续需要高频项来拟合，且由于边缘恰好是水平的和垂直的，功率谱中显示了叠加在其余数据上的垂直线和水平线，这在图 6.11(b)中很明显。对于测试图样，图 6.11(f)中显示了从原频率变换中消除这些线，然后变换回去的结果。图像的中心部分未受影响，但边缘处的不连续已不那么明显，因为用来在边缘处创建台阶函数的许多频率丢失了。边界一侧的图样已被复制到另一侧，并叠加在原始图样上。

6.5　首选方向

图 6.12 显示了周期性结构的另一个例子，即由偏振光观察的磁性材料薄膜上的周期性结构。条纹的方向与材料上用于在胶片上存储信息的磁畴方向相反。这幅图像的频率变换显示了磁畴的间距，但

不是单峰，而是弧线，这些弧线显示了条带方向的变化，原始图像中的条带方向变化很明显，但难以量化。

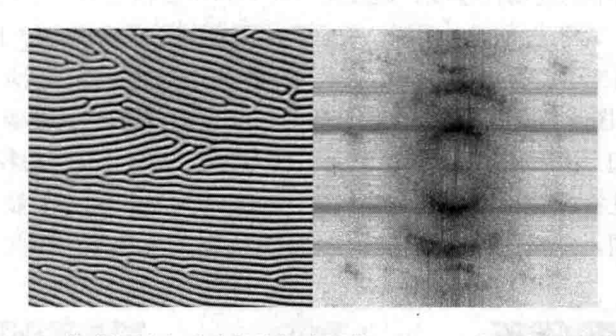


图 6.12　薄膜材料中磁畴的偏振光图像（左）及其频率变换（右）

测量功率谱中弧线的角度范围以及沿这些弧线的亮度（功率）变化，可表征结构中的取向度。即使对于不完全呈周期变化的结构，也可使用与角度有关的整体功率来测量首选方向。这与空间域中执行的自相关运算的结果完全相同，在这种运算中，稍微偏移一幅二值图像后，将偏移图像与之前的图像组合在一起，得到了一个小数值矩阵，但在频率域中完成这种操作的速度更快，且处理灰度值更为容易。

重建不完全对齐的周期性结构，可通过选择频率变换中的整条弧进行。图 6.13 用病毒颗粒对此进行了说明。负染病毒的 TEM 图像暗示在内部的螺旋结构中，但不能清晰地显示出来。在频率变换中，周期性间距和方向变化很明显。测量中心到弧的距离可求出间距，进而由弧的角度范围求出螺旋角。仅使用这些弧线的重新变换结果中，显示了周期性，并使得重建仅限于病毒的内部，进而使得内部的螺旋图案很明显。

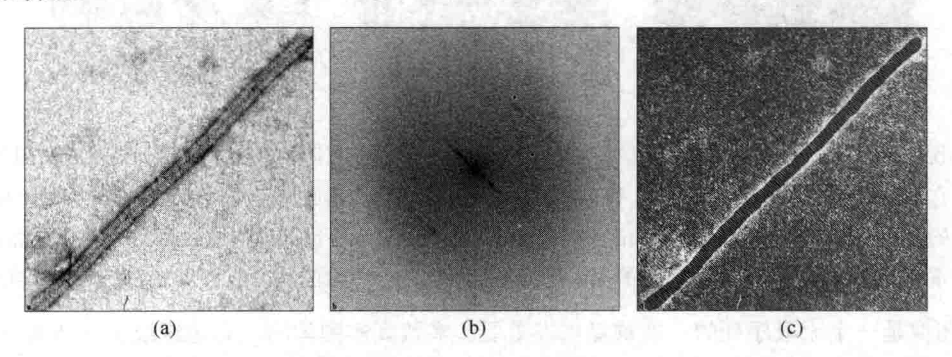


图 6.13　病毒的 TEM 图像：(a)原始图像，内部的螺旋结构难以观察到；(b)图(a)的频率变换，其中病毒的规则重复结构及其方位角变化很明显；(c)仅反变换频率变换中的几个峰，将重建限定在了病毒的内部

傅里叶变换功率谱中由于损失了高频数据，因此图像很模糊。如图 6.14 所示，对于聚焦良好的图像，与频率（半径）相关的平均振幅通常会逐渐下降，但用于在图像中产生清晰边缘的高频信息丢失时，平均振幅会急剧下降。许多自动聚焦算法试图最大化高频，尽管图像中的噪声会对该算法造成负面影响。

图光这种图像缺陷会导致图像失焦，并使某个方向上的清晰度下降。对电子显微镜来说，由于电磁透镜的操作原理，这更是一个突出的问题。即使是熟练的操作者，消除散光也需要花大量的时间，且为校正目的而从图像中识别出它也很困难。在频率变换中，散光会立刻变得明显，因为（在高频处）亮度或功率的放射状下降和非对称性非常明显。

图 6.15 显示了一个例子。这是多晶材料的一幅高分辨率图像。多晶晶格中的规则结构给出了期望

的衍射图样，但垂直方向功率的快速下降表明该图像在散光方向是模糊的。校正散光产生了其功率谱振幅表现出更大均匀放射状下降的一幅图像。

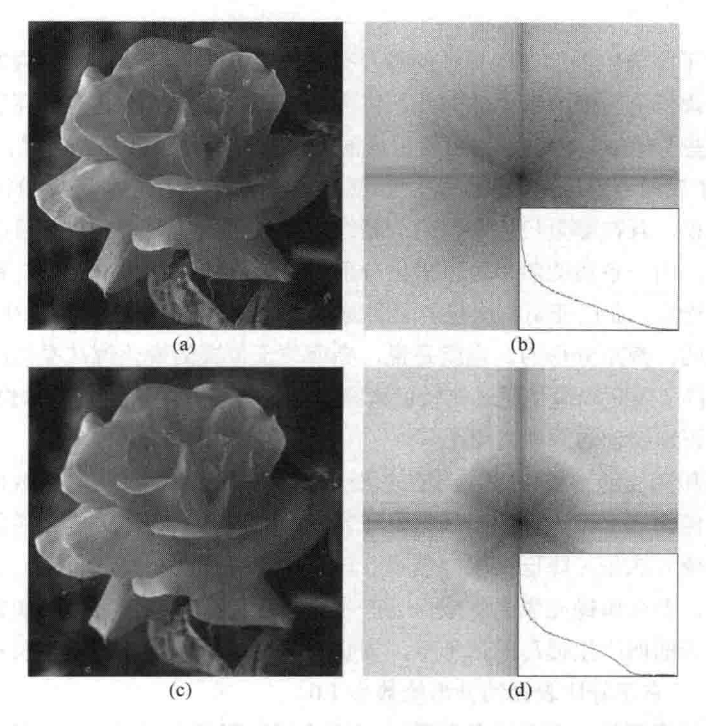


图 6.14 聚焦检测：(a)聚焦图像；(b)亮度值的傅里叶功率谱及平均功率与半径的关系曲线；(c)轻微失焦的图像；(d)傅里叶变换功率谱及平均功率与半径的关系曲线，表明高频处缺失信息

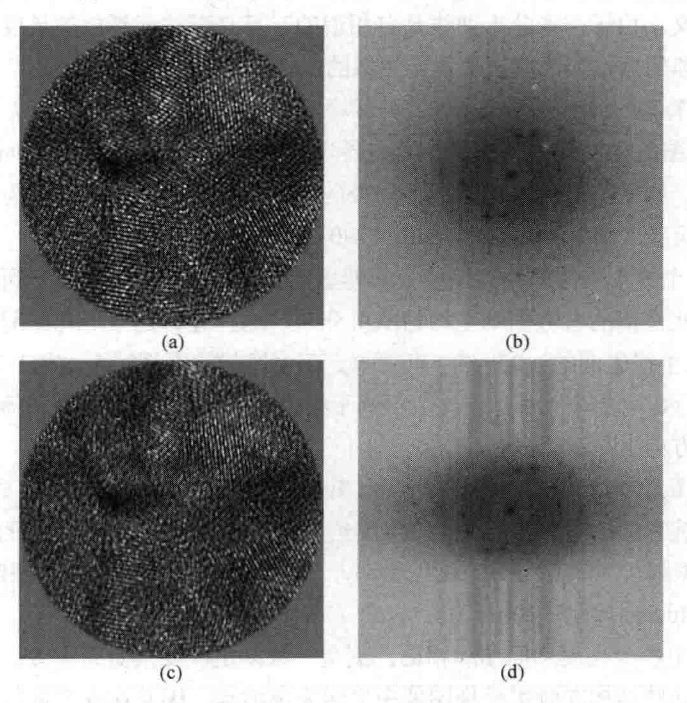


图 6.15 散光对功率谱的影响：(a)无散光图像；(b)图(a)的功率谱，各方向均匀；(c)散光图像；(d)图(c)的功率谱，顶部和底部有信息缺失

6.6 纹理和分形

在功率谱中，除了图像中周期结构导致的峰及有限图像分辨率导致的高频限制外，看起来好像只剩下包含无用信息的背景了。但情形并非如此。许多图像会显示表面反射光的亮度，或显示如表面高程这样的数据。在这些图像中，由功率谱可给出表面的粗糙度或纹理并加以测量。

第 15 章中讨论了表面分形维数的概念。分形表面上具有一个数学上未定义的面积。它大于由不规则表面覆盖的投影面积，且在测量尺度变得更加精细时，该面积会增大。在测量面积与测量工具尺寸之间关系的对数图上，由一条线的斜率可以求出分形维数。许多由磨损、腐蚀、颗粒聚集或断裂造成的自然表面具有这种特性。业已证明，这些表面图像，无论是由漫射光的散射产生的，还是由 SEM 中的二次电子发射产生的，都是分形的。也就是说，亮度关于位置的变化遵从类似的数学关系。此外，如第 5 章中指出的那样（当时测量是通过比较像素亮度值与像素间距离的关系来进行的），分形维数也是表征图像中可视觉识别纹理的一种有用工具。

分形维数是表面粗糙度的一种强大且简洁的表示，它通常与表面历史和所致性质相关。直接测量表面分形并不实际，但物理上可通过求出（粘附在表面上的与其大小相关的）各种气体分子的数量来实现。一种常用的成像方法是，降低维度并测量（该表面与采样面相交产生的）边界线的分形维数。这要么由横断面产生，要么由抛光表面生成孤岛产生。在任何一种情形下，表面越粗糙，交线就越不规则。在抛光平行于表面而产生孤岛的情形下，周长有一个分形维数（待测线长和测量工具长度之间关系的对数图的斜率），它正好比表面的分形维数小 1.0。

对于分形表面的距离图像（亮度代表高程），功率谱显示了所有频率和方向的正弦叠加，并且有一个特定的特征：傅里叶变换中系数的振幅随频率的对数指数地降低，而各项的相位是随机的。这可以定性地理解，按照定义，由于一条分形曲线是自相似的，且具有扩展到更细尺度（或更高频率）的细节。这也意味着高频项的振幅的比例必须是自相似的。指数曲线满足这一标准。

画出振幅（或功率谱，它是振幅的平方）与频率的对数关系图，可生成一条可分析的直线。Feder（1989）已证明，剖面图的分形维数与傅里叶级数展开中各项振幅的指数下降之间，存在一种简单的关系。这种相关使得在二维傅里叶变换图像中使用径向递减振幅来测量粗糙度成为可能，也使得测量该递减的方向变化成为可能（Mitchell & Bonnell，1990；Russ，1990，1994）。

图 6.16 显示了一个分形表面的距离图像及其傅里叶变换功率谱。图 6.16(c)所示振幅与频率间的对数关系为一条直线，相位值的直方图则呈均匀随机分布，这种分布是分形几何所需要的。该图的斜率［该斜率必须在（欧几里得表面的）2.0 与（有效填充三维空间的不规则表面的）2.999…之间］给出表面维数是分形维数 =（6 + 斜率）/2，或对所示例子约为 2.2。这是通过对金属喷砂产生的一种各向同性表面，因此在所有方向上斜率相同。

图 6.17 显示了由机器加工产生的另一个表面，其上具有特定方向的刀痕。对该表面的类似分析［见图 6.17(c)］表明，虽然它仍然是分形的（平均维度约为 2.25），但不是各向同性的。违反各向同性的方法有两种：一是改变振幅与频率对数关系图的斜率，二是改变其与方向有关的截距。第 15 章中将详细介绍分形表面几何（Russ，1994，2001b）。

与测量原始图像中各个亮度剖面曲线相比，测量二维傅里叶变换更为有效。后者可忽略出现的任何周期性结构，因为这些结构在频率变换图像中显示为离散点，因此在求振幅值的整个指数下降时可以忽略它们。换言之，可以研究图像中因数据获取导致的非周期性结构或噪声，并表征表面上混沌但自相似的性质。

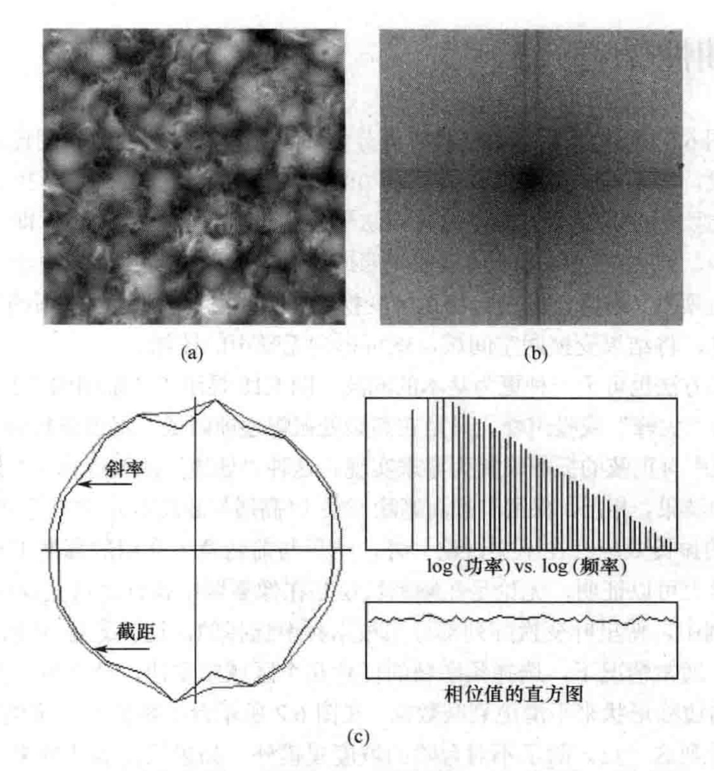


图 6.16　距离图像的分形分析：(a)原始图像（喷砂处理的金属表面）；(b)傅里叶变换功率谱；(c)文中描述的傅里叶变换分析

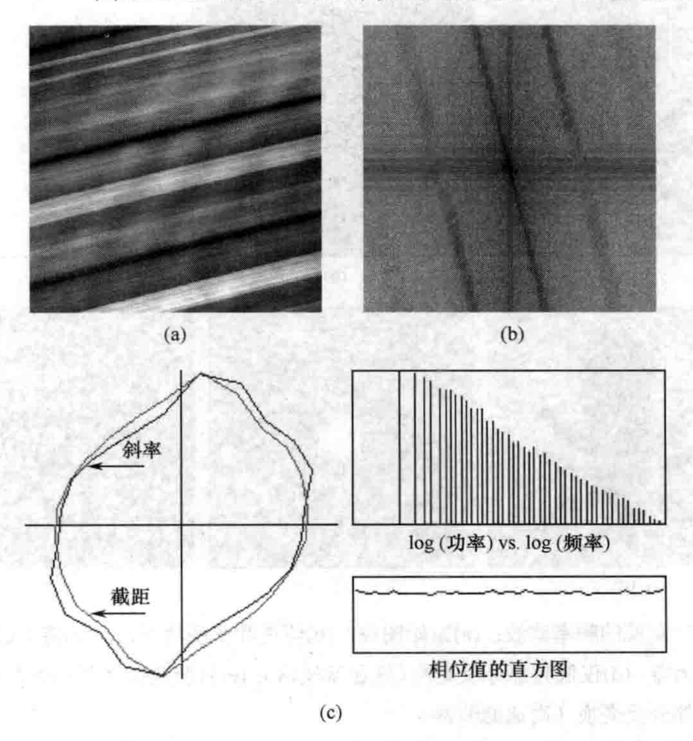


图 6.17　各向异性表面的分形分析：(a)原始图像（机器加工金属表面）；(b)傅里叶变换功率谱；(c)文中描述的傅里叶变换分析

6.7 隔离周期性噪声

如上所述（及图 6.5 所示），频率变换具有可分性和可加性。将两幅原始图像或两个函数的变换加到一起所产生的结果，与两幅原始图像或函数之和的变换结果是相同的。这一思想就为使用减法操作来删除图像中的不想要部分找到了方法。这种方法最常用于删除周期性噪声，即由用于记录或传送图像的设备引入的噪声，或由某些环境干扰（如光照或振动）导致的噪声。下面给出了几个例子。如果确定了与噪声相关的频率（通常直接由自傅里叶变换确定），然后将那些频率项的振幅设为零，则可只保留所需的部分信息。将结果变换回空间域，就可得到无噪声的图像。

删除所选频率的方法也可于一种更为基本的滤波。图 6.18 显示了一幅图像及其傅里叶变换功率谱。功率谱中没有明显的“尖峰”或噪声峰，只是在高频处振幅逐渐降低。保留低频而删除高频，可通过将一个选定频率之上的所有正弦的振幅设置为零来实现。这种“低通”滤波运算（即通过或保留低频）产生了图 6.18(d)所示的结果。相反，保留高频而删除低频（“高通”滤波器），产生了图 6.18(f)所示的结果。

除了边界周围的振铃效应（在下文讨论）外，结果与前两章中介绍的高斯平滑及拉普拉斯锐化的结果完全一样。数学上可以证明，无论是在频率域还是在像素域中执行，这些操作的结果完全相同。

在滤波这一示例中，傅里叶变换阵列部分是根据频率选择的，这就是这些滤波器通常称为低通和高通滤波器的原因。通常情况下，选择频率域的任意几个区域来重建，会产生“振铃”缺陷，除非仔细调整滤波器区域的边缘形状来平滑地衰减数据。在图 6.2 所示台阶函数的一维例子及图 6.6 所示的对应二维例子中，可看到这一点。除了不对台阶的坡度建模外，如果只用前几项来重建，那么在边缘附近会出现振荡，这通常称为振铃。在图 6.18(d)和图 6.18(f)所示的重建图像中可看到振铃现象。

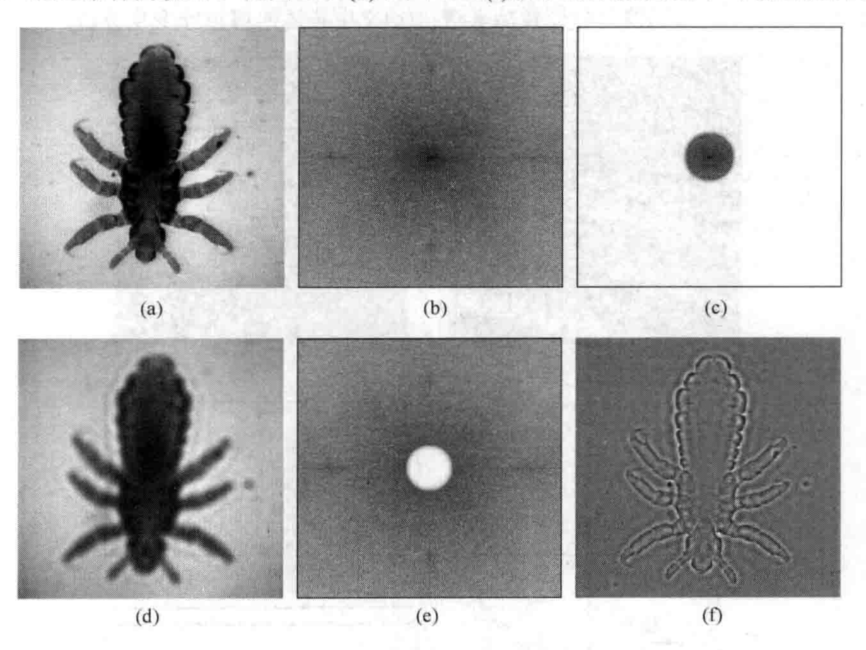


图 6.18　使用傅里叶变换的频率滤波：(a)原始图像；(b)傅里叶变换功率谱；(c)将所选值之上的所有频率的振幅降为零；(d)仅低频部分反变换（低通滤波器）；(e)将所选值之下的所有频率的振幅降为零；(f)仅高频部分反变换（高通滤波器）

调整滤波器边缘的形状，可防止在明显的不连续处出现振铃。这种行为就是众所周知的一维滤波（常用于数字信号处理中）。常用几种不同的形状。在某个指定的距离上，使用一个加权函数来降低滤

波器振幅。最简单的函数是线性内插（也称为 Parzen 窗函数）。使用抛物线或余弦函数（在上下文中分别称为 Welch 和 Hanning 窗函数），可以得到更好的结果。最精细的滤波器形状并不会降为零，但在截止点之外会延伸很长的尾巴。这样的一种形状是高斯滤波器。图 6.19 显示了几个此类形状。

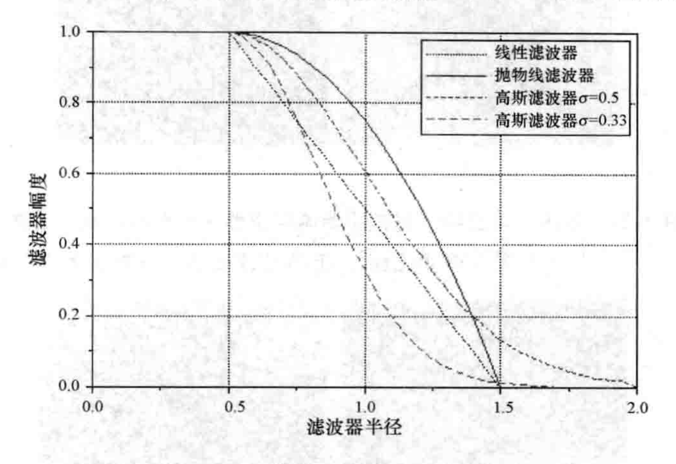


图 6.19　一些常见的滤波器边缘曲线

这些应用中经常用到的另一个滤波器形状是巴特沃思滤波器，其振幅可以写为

$$H = \frac{1}{\left[1 + C\left(R / R_0\right)^{2n}\right]} \tag{6.8}$$

式中，R 是到滤波器中心（功率谱图像的中心或零频率点）的距离，R_0 是标称滤波器截止值。常数 C 一般设为 1.0 或 0.414；该值定义了在 $R = R_0$ 点处滤波器的振幅，要么为 50%，要么为 $1/\sqrt{2}$ 。整数 n 是滤波器的阶数，最常用的取值是 1 或 2。图 6.20 比较了几种巴特沃思低通滤波器（衰减高频的滤波器）的轮廓。带负值 n 的相反形状通过高频并阻隔低频。图 6.21 显示了调整频率截止形状对图 6.18 所示图像进行高通和低通滤波质量的影响。

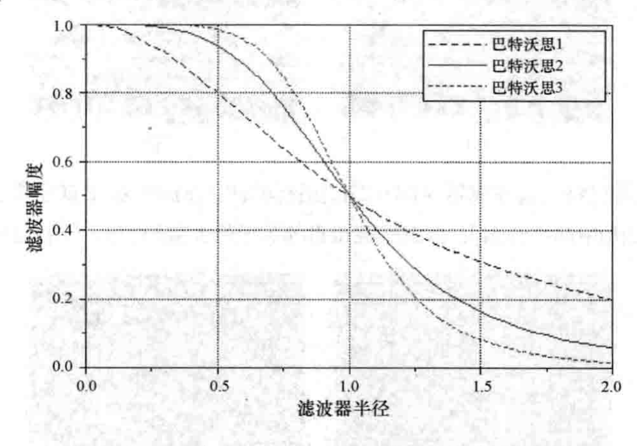
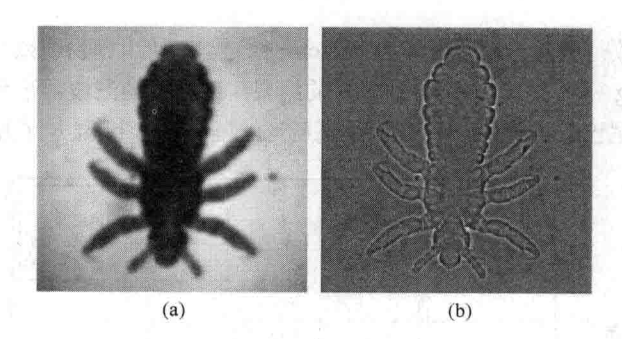


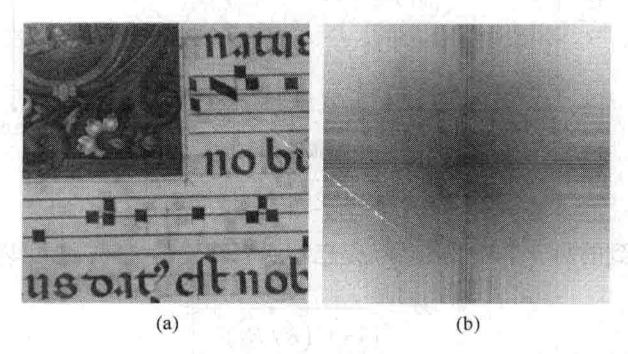
图 6.20　一阶、二阶及三阶巴特沃思滤波器曲线

图 6.22 显示了一幅具有细节和噪声的图像及该图像的频率变换。对频率域图像应用半径为 10 和 25 像素的二阶巴特沃思低通滤波器，平滑了噪声并模糊了一些高频细节（见图 6.23），而应用具有同样半径的巴特沃思高通滤波器则强调了边缘并降低了大（低频）区域间的对比度（见图 6.24）。所有这些滤波器都作为乘性掩模应用到振幅值，而相位值保持不变。调整滤波器截止的半径，可控制逆变换后图像中的频率范围，就像在空间域或像素域中执行等效的过程时，改变卷积核的大小来改变平滑度或锐化度那样。



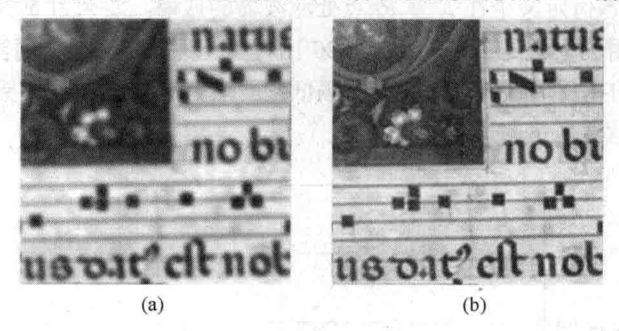
(a) (b)

图 6.21 使用二阶巴特沃思滤波器调整频率截止形状：(a)低通滤波器［与图 6.18(d)比较］；(b)高通滤波器［与图 6.18(f)比较］



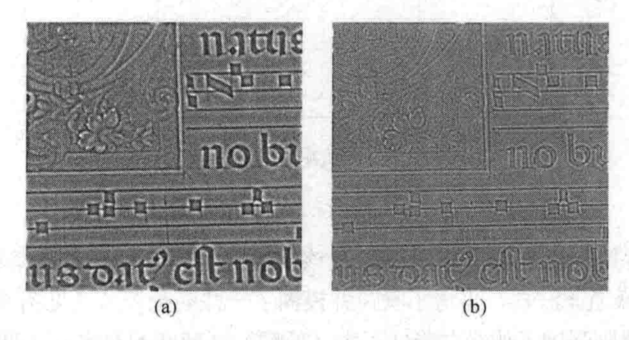
(a) (b)

图 6.22 图 6.23 和图 6.24 中用于滤波的图像及其变换：(a)原始图像；(b)亮度数据的傅里叶变换功率谱（红色圆圈表示巴特沃思滤波器所用 50%截止点的半径）



(a) (b)

图 6.23 使用低通巴特沃思滤波器对图 6.22(a)进行滤波：(a)50%截止直径等于最大值的 10%［图 6.22(b)中的小圆圈］；(b)50%截止直径等于最大值的 20%［图 6.22(b)中的大圆圈］



(a) (b)

图 6.24 使用高通巴特沃思滤波器对图 6.22(a)进行滤波：(a)50%截止直径等于最大值的 10%［图 6.22(b)中的小圆圈］；(b)50%截止直径等于最大值的 20%［图 6.22(b)中的大圆圈］

振幅减少到零是没有必要的。有时会将下限设置为一个小数，以便高（或低）频不会完全衰减。也可使用大于 1 的值。低频值为一个小值（如 0.5）而高频值大于 1 的高频强调滤波器，可应用到其亮度值之前已被转换为对数的图像。使用对数亮度数据很重要，因为它会使得光照和反射的乘积效果变为各加项。对这些值应用高通滤波器，可抑制光照中的（大）变化，并使得反射中的（较小）变化更为明显，进而锐化细节的外观。第 3 章中的图 3.10 说明了该结果。这个同态滤波器的物理原理是，就如多数此类滤波器的作用那样，分离图像中的光照分量和反射分量，且确实可以改善许多实际图像的外观。

也可以选择中心不在原点处的傅里叶变换图像的区域。图 6.25 中选取了某个方向范围内的中间频率值，并显示了重建结果。此类滤波可用于从图像中选取方向性信息。变换中的频率分量方向垂直于空间域中对应周期线的方向。删除所选频率与方向的方法相同，如下所示。该图还说明了傅里叶变换图像的基本特征：傅里叶变换图像中的位置，标识了部分或全部空间域图像中的周期和方向信息。如所示的那样，离心区域的正确形状是环形与楔形（其截止对应于频率和角度）的组合，而非更易于构造的圆形或矩形，以及上面所示的需要同类形状的边缘。

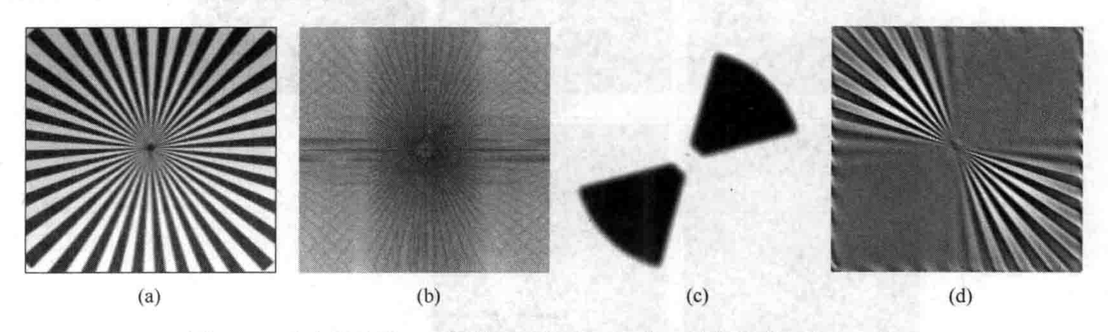


图 6.25　方向性滤波：(a)原始测试图案；(b)傅里叶变换功率谱；(c)保留某个方向范围中各项的掩模；(d)掩模保留频率的反变换

6.8　选择性掩模和滤波器

原始图像中的周期性噪声位置隔离为傅里叶变换图像中的几个点后，删除相关的频率项，就可删除这些噪声。将所选频率项的振幅减小为零，滤波器就可删除这些频率与方向。滤波器并不改变相位信息。

指定和应用这种降低振幅的方法有多种。有时以数值方式指定频率和方向的范围较为实际，但更为可取的方法是，使用傅里叶变换图像的振幅谱或功率谱。手动或自动选取傅里叶变换图像中的区域，允许我们为消除目的而选取功率谱中对应于周期信息的各个峰。此处所用的术语“滤波器”和“掩模”本质上是同义的（在某些教材和程序中，掩模被理解为仅由 0 值或 1 值组成，而滤波器则有连续的衰减值；由于选择用来删除所选频率的掩模通常仍具有截止的边缘形状，因此此处交替使用这两个术语）。

在某些情形下，可以数学上构建一个滤波器并将其直接用于存储的变换，但它通常创建为一个像素阵列，该阵列的大小与变换的大小相同，其中包含有与 FFT 数据相乘的值。这些值可能只是删除或保留原始各项的“开”或“关”，但很多情形下它们具有中间值，即除范围 0～1 内的值外，它们还可由图像中所用的值域 0～255 来表示。这意味着掩模或滤波器本身就可当做一幅图像来处理，而且可对其及其功率谱应用所有的图像处理工具，如平滑边缘的高斯滤波器。

替代组合使用弧与径向线，对于分隔开的噪声峰（通常称为噪声“尖峰”），当它们由一些小点构成时，用小圆圈来定义区域是很方便的，且通常是可接受的。重要的是，要使用平滑函数来修改区域

的边缘。当过渡发生在只有几个像素的距离上时，上面显示的各种过渡曲线间的差异就不重要了，此时最可取的方法是应用一个简单的高斯平滑。

图 6.26 说明了从傅里叶变换中删除相应频率，进而从图像中删除周期性噪声的过程。在原始图像中，选择一个大小为 2 的整数次幂的方形区域，并计算其 FFT。两个暗的“尖峰”（每个在功率谱中出现两次）对应于图像主噪声的方向和半径。删除这两个尖峰后，就消除了噪声。注意，所处理区域的四周仍可看到噪声（且噪声稍向处理区域跨越了一些）。还要注意，所有细线和细节都被保留，因为它们是由不同的频率组成的。

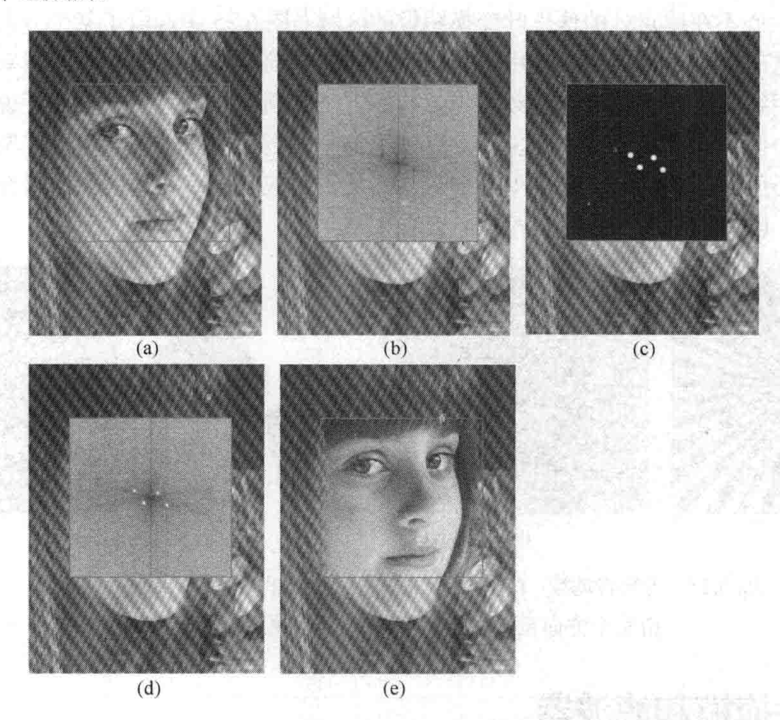


图 6.26 删除周期性噪声：(a)圈出待处理区域的原始图像；(b)傅里叶变换功率谱；(c)对应于噪声尖峰的掩模；(d)删除噪声尖峰后的结果；(e)反变换结果

通常会存在多个尖峰，只有先检查功率谱，才可确定这些点的存在和确切位置。图 6.27 显示了杂志上的一幅半色调印刷图像。印刷工艺中所用的半色调网屏产生了这种图案。在图像的频率变换中，这种规则图案显示为明显的窄峰或尖峰。将这些位置的振幅设置为零（但不改变相位信息），可删除这些尖峰。这样，进行反变换时图像中就无噪声了。

图 6.26 所示例子依赖于人来观察傅里叶变换图像中的尖峰，识别出尖峰是周期性噪声产生的原因，并人为选取它们来生成滤波器。许多情形下，可由傅里叶变换功率谱自动地构造一个合适的图像滤波器。指导原则是，删除功率谱中狭窄且明显高于局部背景的尖峰。如果像处理平常的空间域灰度图像那样来处理傅里叶变换功率谱图像，那么使用如顶帽这样的排序滤波器（见第 4 章）就可自动地删除这些尖峰。这是图 6.27 中使用的方法；滤波器会找到这些尖峰，包括位于原点的点，但在滤波器中必须删除该点，并在变换中保留，因为它表示了图像的平均亮度。

原始图像是彩色印刷图像时，情况会变得稍微复杂一些。如第 3 章及图 6.28 所示，彩色半色调印刷图像通常是使用青色、品红色、黄色和黑色油墨生成的，每种颜色都有一个方向不同的半色调网屏。所需过程是分离颜色通道，处理每个通道，然后合并它们得到最终结果。图中，对应于颜色，显示了每个通道的功率谱的叠加，可看到不同的网屏方向。对每个通道的傅里叶变换功率谱应用一个顶帽滤

波器，选取了各个噪声尖峰，然后消除这些尖峰。再后反变换各个通道。重新合并这些冗余的颜色通道图像，最终生成了显示的结果。当然，这种方法取决于半色调点的规则周期间隔，它不适用于使用不规则或随机间隔彩色点的喷墨打印机打印的图像。

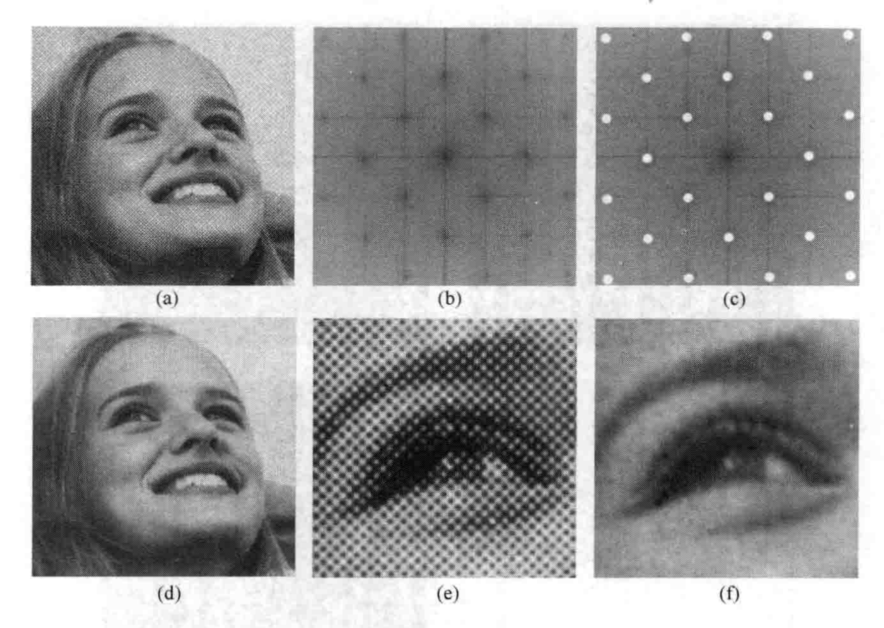


图 6.27 移除复杂的周期性噪声：(a)原始半色调印刷图像；(b)傅里叶变换功率谱；(c)从FFT 中移除尖峰；(d)变换回去的结果；(e)图(a)的放大细节；(f)图(d)的放大细节

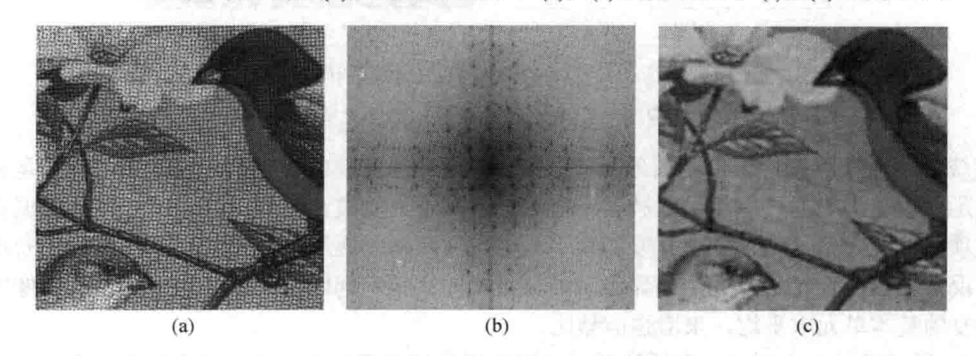


图 6.28 删除彩色图像中的半色调：(a)原始图像（以 CMYK 方式印刷的邮票放大局部）；(b)每个颜色通道的傅里叶变换功率谱叠加后的着色效果；(c)从每个通道中删除噪声并重新合并这些颜色通道的结果。当然，由于本书是使用 CMYK 和半色调印刷的，因此打印过程中引入的另一组频率和间距限制了我们能看到结果中细节的能力

6.9 周期性信息的选择

在有些类型的图像中，有用的是周期性信息，而必须抑制非周期噪声。用来定位周期尖峰、构建滤波器、平滑滤波器边缘等的方法不变。唯一的区别是，滤波器的意义改变了，且滤波器的值在乘以变换振幅之前已被取反，因此保留了对应于周期性信息的尖峰，而消除了对应于随机噪声的背景。

图 6.29 显示了结晶陶瓷（富铝红柱石）的一幅高分辨率 TEM 晶格图像。我们可以看到晶格的二维周期性，但它叠加在一个变化和带噪的背景上，这改变了局部对比度，使得观察这种材料中相当复

杂的晶胞细节更加困难。傅里叶变换中有表示周期性结构的尖峰。如前所述，这幅图像本质上与可以用 TEM 逼真地记录来将标本的衍射图样投影到相机平面图是一样的。当然，因为相位信息已经丢失，从拍摄的衍射图样来对空间域图像进行反变换是不可能的。

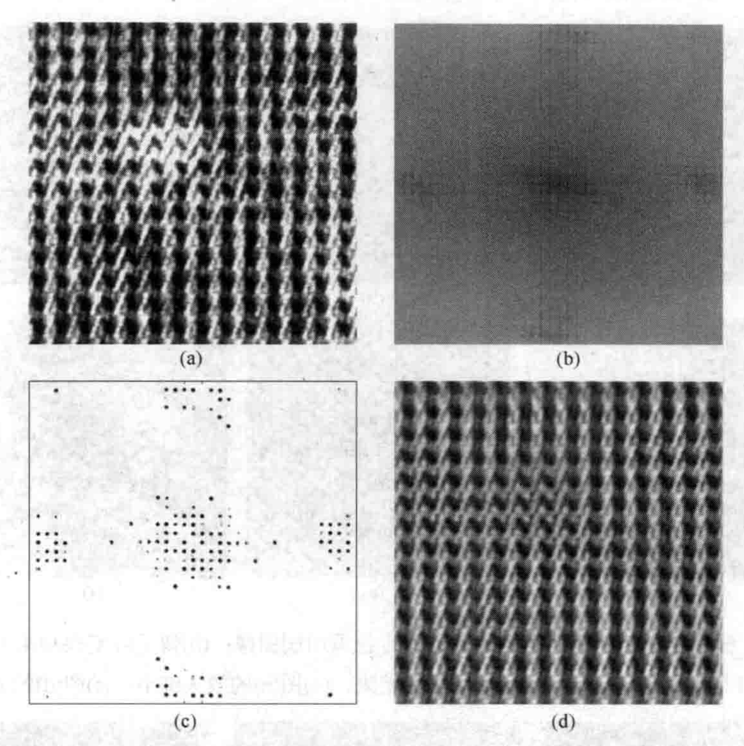


图 6.29　透射电子显微晶格图像：(a)原始图像；(b)傅里叶变换功率谱；(c)对图(b)应用顶帽滤波器来构建一个滤波器的结果；(d)仅对所选周期信息反变换的结果

要围绕每个周期点选择一个小圆形区域，需要构建一个滤波器或掩模。这样做的方法是使用尺寸刚好大到足以覆盖尖峰的一个顶帽滤波器。在这种情况下，滤波器用于保持尖峰中各项的振幅，而将所有其他项降低至零，这就在整个亮度的短小（高频）颗料和逐渐（低频）变化中删除了随机或非周期噪声。反变换该图像，会生成清楚显示了晶格结构的一幅空间图像。这相当于对原始图像中出现的几百个重复的基本单元求平均，来增强信噪比。

图 6.30 显示了另一个例子。在原始图像（肌肉原纤维截面）中，由于存在噪声，因此完全不可能分辨出周期性结构。在有些孤立的位置，可看出一些原纤维有较规律的间隔和排列，但我们并不能很容易地透过噪声和变化来找到这种规律。傅里叶变换显示了底层结构的尖峰（还揭示了有限的图像分辨率）。通过对功率谱图像应用一个顶帽滤波器，可只选取这些尖峰（并保留原始相位信息），并产生了所示的结果。反变换后的图像清楚地显示了肌原纤维结构的六重对称。图中局部放大部分可见粗丝和细丝。而在原始图像中则无法清楚地看到细丝。

使用此类滤波来提取周期性结构时，一定要小心，原因是构建一个掩模也可能会删除图像中的真实信息而保留缺陷和噪声。在六重对称的功率谱中选取这些点，可确保经滤波并反变换后的空间图像中显示这种类型的结构。这意味着关键的一步是在傅里叶变换中识别和选取尖峰。

处理功率谱图像来构建掩模还有许多其他方法。图 6.31 显示了一个例子，由于各尖峰靠得太近，因此使用顶帽滤波器来选取它们很困难。对该功率谱图像的一个副本应用中值滤波器，删除了尖峰，生成了一幅背景图像。从原始图像中减去这幅背景图像（此时是适当的，因为功率谱图像显示了对数振幅）调匀了对比度，进而使用阈值处理方法可选取所有的尖峰。将经阈值处理后的图像作为掩模，

并反变换该图像，可得到这一平均的结果。在这种情况下，处理的是原始图像的亮度通道，滤波后的结果会回插到亮度通道，而保持颜色信息不变。

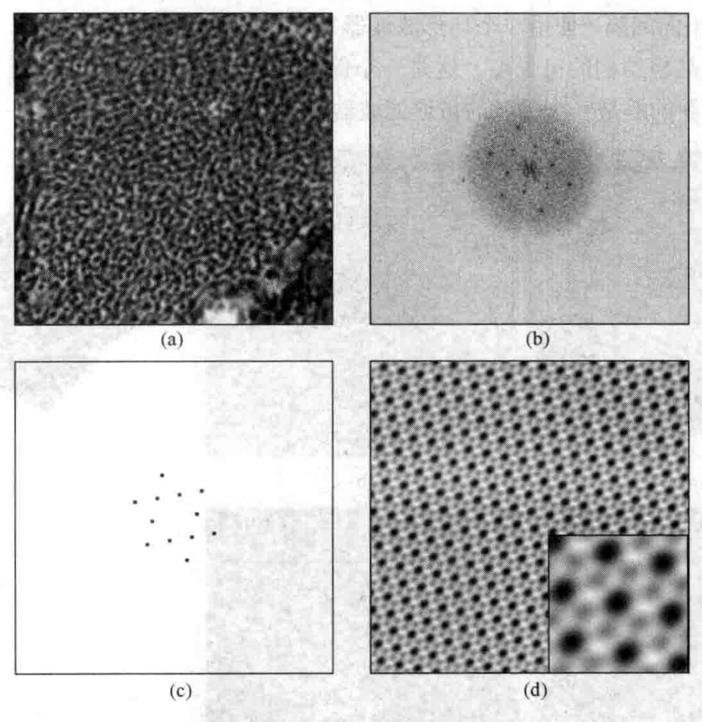


图 6.30 肌肉原纤维截面中的周期性：(a)原始图像；(b)傅里叶变换功率谱；(c)由顶帽滤波器找到的尖峰；(d)细节放大的重建图像

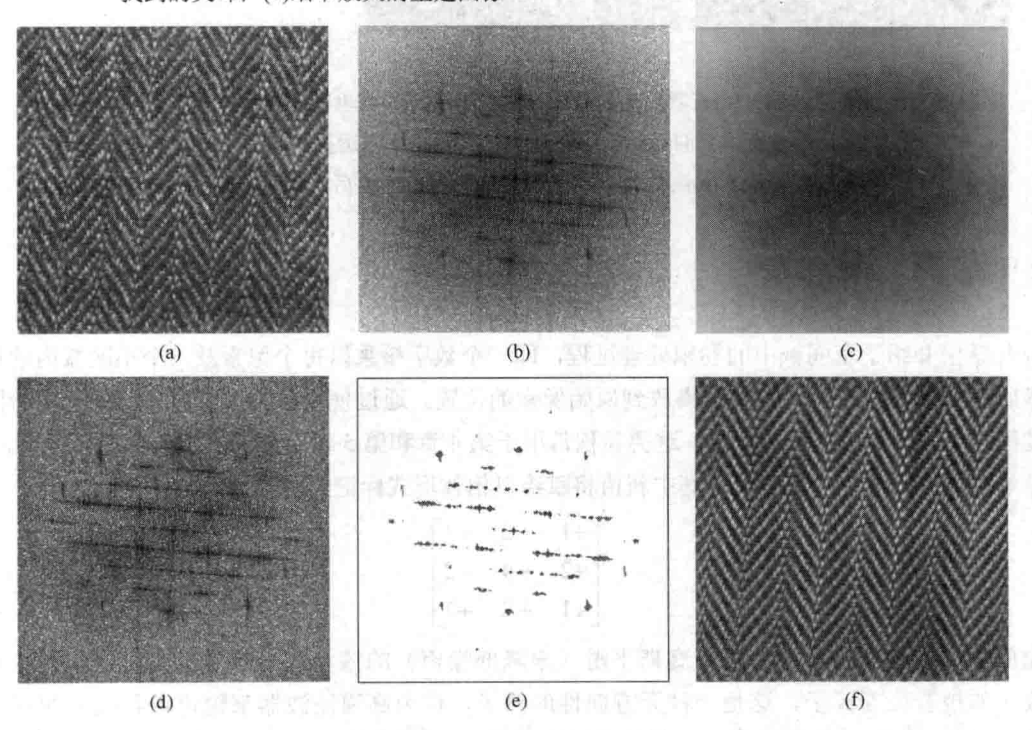


图 6.31 周期性结构的重建：(a)原始图像；(b)傅里叶变换功率谱；(c)对图(b)应用中值滤波器后的结果；(d)图(b)减去图(c)后的结果；(e)图(d)经阈值处理后生成一个掩模；(f)反变换后的结果

也可构造一个滤波器来选取某个较窄范围的频率。环形滤波器可增强所选的周期结构。图 6.32 显示了轮胎制造中所用石墨化碳颗粒的高分辨 TEM 图像。功率谱与衍射图样相同，并以点环的形式显示了约 3.5 Å 的主原子平面间隔。使用一个环形滤波器只选择该间隔，然后反变换，再将原子位置添加到原始图像上，增强了晶格结构的可见度。这是一个保留了选定频率范围的“带通”滤波器的例子（在空间域中，高斯滤波器的差是一个等效的带通滤波器）。

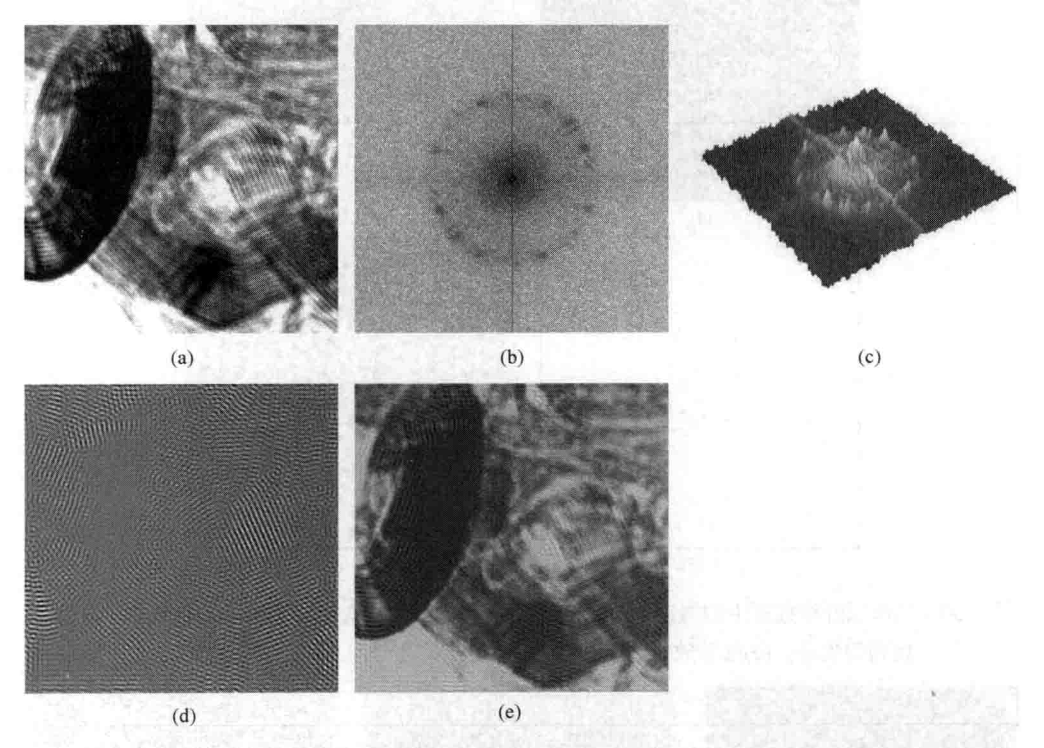


图 6.32 带通滤波器：(a)石墨化碳的原始 TEM 图像；(b)傅里叶变换功率谱，显示了石墨中对应于 3.5 埃平面间距的不完整圆环；(c)功率谱透视图强调了圆环尖峰；(d)圆环点的反变换；(e)图(d)插入原始图像中的红色通道后，增强了原子晶格的可视性

6.10 卷积

前几章中介绍了空间域中的卷积处理过程，即一个数字核乘以每个像素及一个小区域内的相邻像素，然后对结果求和，再将求和结果放到原始像素的位置。通过使用原始像素值并产生一幅新图像，将该过程应用到图像中的所有像素。这类卷积常用于第 4 章和第 5 章中介绍的平滑和求导运算。例如，一个简单的平滑核可包含以下值，其中新值将赋给以粗体形式标记的位置：

$$\begin{bmatrix} +1 & +2 & +1 \\ +2 & \mathbf{+4} & +2 \\ +1 & +2 & +1 \end{bmatrix}$$

空间域中的核有多种，如用于高斯平滑（来降低噪声）的核，取一阶导数（如定位边缘）或二阶导数（如拉普拉斯算子，这是一种无方向性的算子，作为高通滤波器来锐化点和线）的核。它们通常以一组整数的形式出现，尽管使用浮点数可得到更高的精度，但取整形式时可认为存在一个除数（通常等于这些数之和或最大的正数）来归一化结果。包含负值的数字核会产生相反的结果，此

时结果中要添加一个偏移值（对像素值域为 0～255 的图像而言，这个偏移值为 128，即中间灰度值）；或采用自动标定方法，将算术结果匹配到值域 0～255。有些算子会明显大于上面给出的 3×3 算子（此时会将一个很大区域中的相邻像素的加权和加到一起），且不必是方形的。

使用大核会很耗时。图 6.33 以图形方式说明了核的单次放置过程。即使使用快速的现代计算机，并优化处理代码以最有效的顺序执行加法和乘法，用大小为 25×25 的核对一幅千万像素的图像执行这一操作，也需要大量的计算，更不用说更大的核和图像。尽管使用专用硬件（如流水线阵列处理器）可加速这一进程，但费用高昂。可使用一种类似的方法来加速傅里叶变换的计算，但此处的关注点是算法而不是它们的实现。

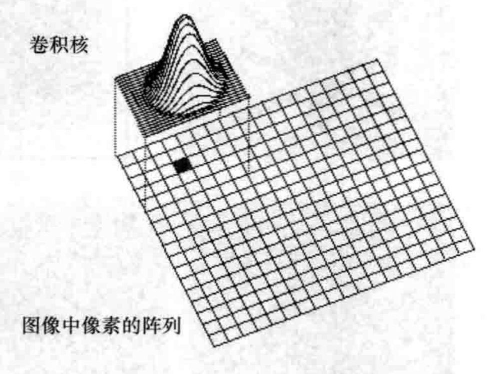


图 6.33　对空间域中的一幅图像应用卷积核的说明

对于任何基于计算机的系统，核大小的增加，最终会达到在傅里叶域中更有效执行这一运算的临界点。从空间域到频率域的 FFT 变换及反变换所需的时间，要比执行卷积操作的速度快很多。如果存在其他原因需要将图像变换到频率域，那么即使是很小的核，计算效率也要高很多。

这是因为等效于空间域卷积（需要很多次乘法和加法）的这种运算，在傅里叶变换中是每个振幅值与核变换中相应值的一次相乘，这样就变成了一个滤波器或掩模。就像存储核那样，核（即滤波器）变换也可事先存储。空间域中的卷积完全等效于频率域中的相乘。使用前面的符号表示，即图像是函数 $f(x,y)$ 而核是 $g(x,y)$，在图像中的每个位置放置核并与之相乘的卷积运算为

$$g(x,y)\otimes f(x,y)=\iint f(\alpha,\beta)\cdot g(x-\alpha,y-\beta)\mathrm{d}\alpha\mathrm{d}\beta \tag{6.9}$$

式中，α 和 β 是积分哑变量，其范围跨越整幅图像，而符号⊗表示卷积。若图像 $f(x,y)$ 和滤波器 $g(x,y)$ 的傅里叶变换分别是 $F(u,v)$ 和 $G(u,v)$，则傅里叶域中的卷积运算就是简单的逐点相乘，或

$$g(x,y)\otimes f(x,y)\Leftrightarrow G(u,v)\quad F(u,v) \tag{6.10}$$

两种运算间有一些不同。由于边缘像素不存在相邻像素，因此空间域中常用的核会对边缘像素应用特殊规则。但将图像变换到频率域中后，会假设图像在边缘处是环绕的，以便左边缘与右边缘是连续的，顶部边缘与底部边缘也是连续的。频率域中相乘方式的卷积，等效于将核应用到空间图像时，以这种相同的环绕方式寻址像素，且通常在边缘处会产生一些人为缺陷。此问题最常见的解决方案是，在一幅较大的图像［图像中的边缘要么填充图像的平均亮度值，要么由边缘值来内插（即上面给出的“填充”）］中嵌入感兴趣的图像。

图 6.34 显示了对于同一个平滑核，空间域中的卷积结果与频率域中相乘结果。这个核（即标准差为 2.0 像素的一个高斯滤波器）显示为一个灰度值阵列，连同显示了其变换。在空间域中将该核应用到图像，生成了例中所示的结果。该核的变换与图像的变换相乘，生成了显示出了功率谱的频率域图像。反变换该图像，就生成了与空间域操作结果相同的结果。

注意，频率域相乘与空间域邻域处理等效这一结果，仅适用于卷积滤波器（也称为“线性”滤波器）。其他邻域运算，如基于排序滤波（保存邻域中最亮、最暗或中间亮度值）和直方图修正（如局部均衡）都是非线性的，因此在频率域中不存在等效操作。

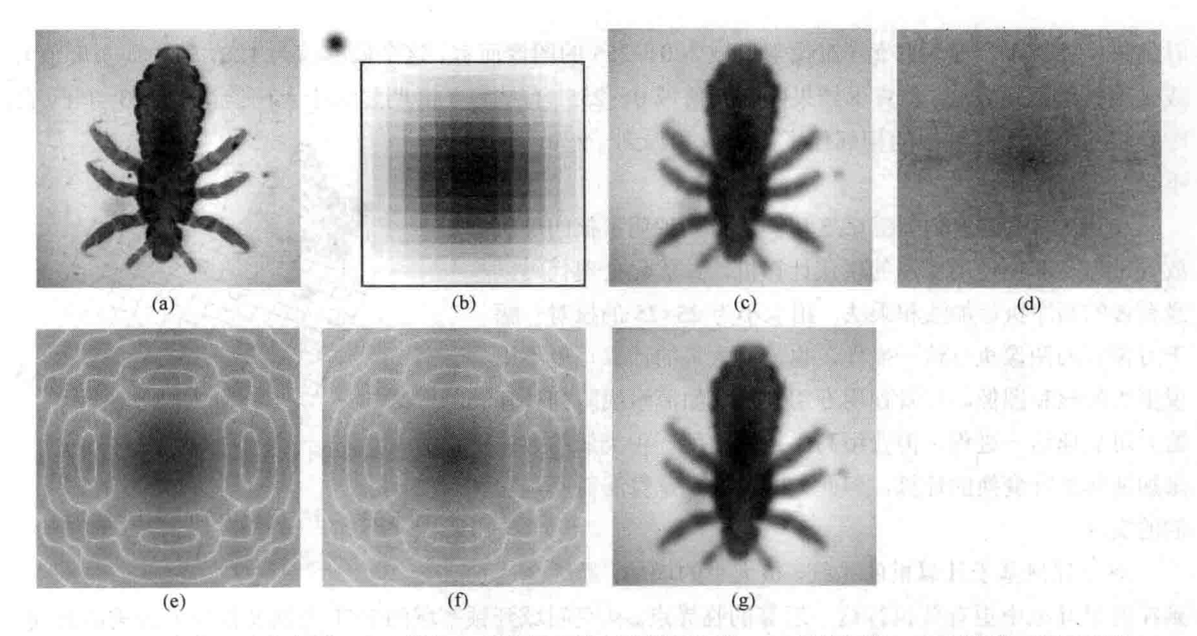


图 6.34　在空间域中应用一个大核与在频率域中卷积的结果比较：(a)原始图像；(b)平滑核（标准差为 2.0 像素的高斯核），已放大到显示出了像素细节；(c)空间中与核卷积生成的平滑图像；(d)图(a)的傅里叶变换；(e)图(b)的傅里叶变换；(f)图(d)与图(e)的积；(g)图(f)的反变换，它与图(c)相同

6.11　去卷积

卷积也可以用来作为理解成像系统如何改变图像或退化图像的一种工具。例如，镜头缺陷导致的模糊可描述为函数 $H(u, v)$乘以图像的频率变换，如图 6.35 所示。物理光学操作通常是在频率域中建模。

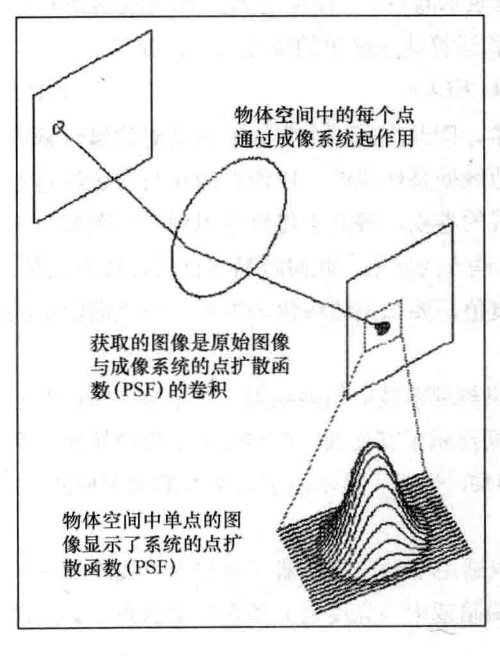


图 6.35　系统特性引入一个点扩散函数（PSF）到已获取的图像

有时，我们可确定系统中每部分的单独特性，但通常并不能确定。在许多情形下，求出系统的点扩散函数（物面中完美点在像面中模糊的程度）后，可用于锐化图像，即删除一些模糊。这称为去卷积，它是通过除以 $H(u, v)$即点扩散图像的变换来完成的。

光学工程师请注意：点扩散函数或 PSF 是在空间域图像中测量的一个理想点的扩展。常用于表征光学系统分辨率和性能调制传输函数或 MTF，则是点扩散函数的傅里叶变换的绝对振幅。高斯形状常用做 PSF 的一个方便模型的原因之一是，其傅里叶变换也是一个高斯形状；原因之二是，几种迭代去卷积技术（如 Van Cittert 法）需要一个像高斯形状那样对称的、中心加权的 PSF；原因之三是，统计学中的中心极限定理，即只要出现大量自变量，那么其组合效果就趋于高斯形状，所以在由光照、标本或物体、光学、照相机和电子学组成的实际系统中，高斯形状通常是一个有用的近似。但在无像差和衍射效应下，散焦模糊的 PSF 呈圆盘状。

图 6.36 说明了卷积和去卷积间的关系。在图 6.36(a)中，一个点扩散函数与一幅图像卷积，产生模糊的结果。

其中*表示卷积运算，它是复算术相乘操作。图 6.36(b)说明了用同样的点扩散函数对模糊图像去卷积的结果（其中÷表示去卷积，它需要复数除法）。这种方法在一定程度上恢复了模糊的图像。但要注意的是，去卷积后图像中的噪声增加了，因此并非所有的原始分辨率都可以恢复（在模糊过程中会丢失一些数据），且在图像边缘会出现人为缺陷。这些都是下面将要处理的去卷积特征。

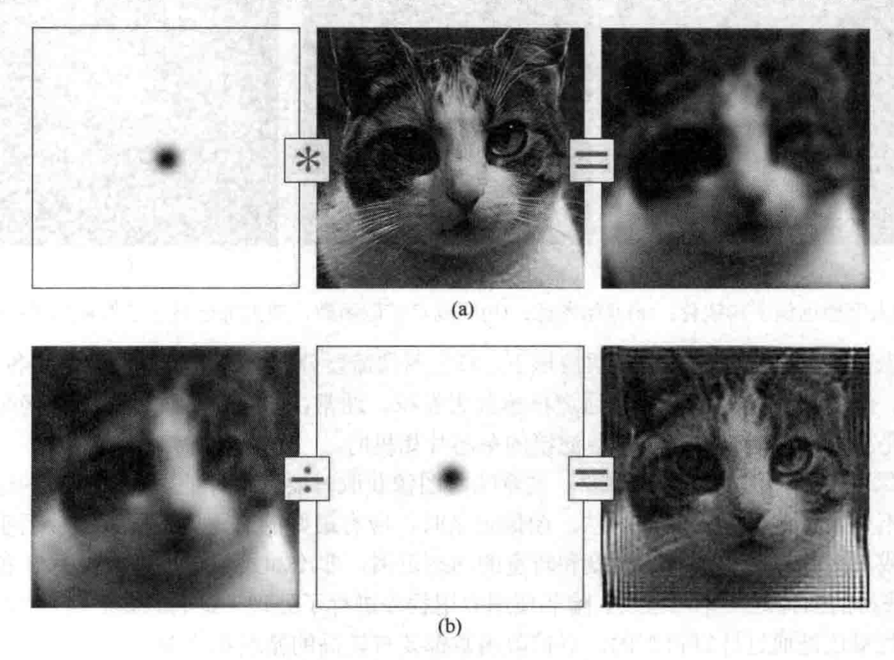


图 6.36　(a)卷积过程和(b)去卷积过程

由成像系统缺陷所致散焦校正的一个知名示例是哈勃太空望远镜。最初部署时，由于哈勃太空望远镜中有一个未正确配置的主反光镜，因此所得图像不能正确聚焦。后来，为补偿多数缺陷，在成像系统中安装了一个补偿光学设备（一个改后的副镜），但即使是使用有问题的望远镜，由已知点扩展函数（或 PSF）来卷积，也可得到高分辨率的结果。但校正后的光学包恢复了望远镜的大量聚光能力，因此由不正确镜面曲率导致的问题得到了大大降低。

在这个特例中，通过对镜面的测量数据，可计算出 PSF。但在许多天文情形下这一问题可简化，因为通过检查单颗星（从地球上看，该星可视为一个点）的图像可测量出点扩散函数。对于图像，这需要相当大的灰度深度，但在天文学中，冷却式相机通常至少能传递 14 位的信息（16000 个线性亮度阶），这对于获得足够精度来进行卷积操作是十分重要的。

图 6.37 显示了这种方法的结果。原始图像的傅里叶变换除以测 PSF 的傅里叶变换，然后反变后所得结果，使得模糊的原始图像的清晰度大大增加。与增大边缘对比度且使得图像看起来更清晰的操作如反锐化掩模（一种高通滤波器）不同的是，去卷积会恢复额外的分辨率并揭示原始图像中不可见的模糊细节。

对于去卷积，方法是使用散焦测试图样的复频率域图像除以点扩散函数的复频率域图像（如图 6.36 所示）。这是复数除法，其执行方法是除以振幅值并减去相位值。该处理的一个问题是，除以非常小的数值会导致数值溢出问题，且对称且良态点扩散函数的傅里叶变换通常包含零值。解决这一问题的常用做法是，使用切趾法将除法运算限制到复变换中不会引起溢出的那些像素上，或让分母加上一个小常数。下面将介绍这两种做法。

对光学散焦、运动、相机芯片中晶体管的大小和排列方式、电子设备带宽不足或其他原因所致的模糊图像，采用去卷积法来恢复并不科学。不存在数学上的最优过程，也不存在适用于所有情形的真理。所用的实用技术均是在许多假设条件下开发的，应用到许多真实情形时明显改进了图像的质量。

尽管结果不错，但它是不同限制条件的折中，其中一个限制条件是计算所需的时间。要成功地应用该方法，要求图像中的随机噪声尽可能少，且计算精度要足够高。

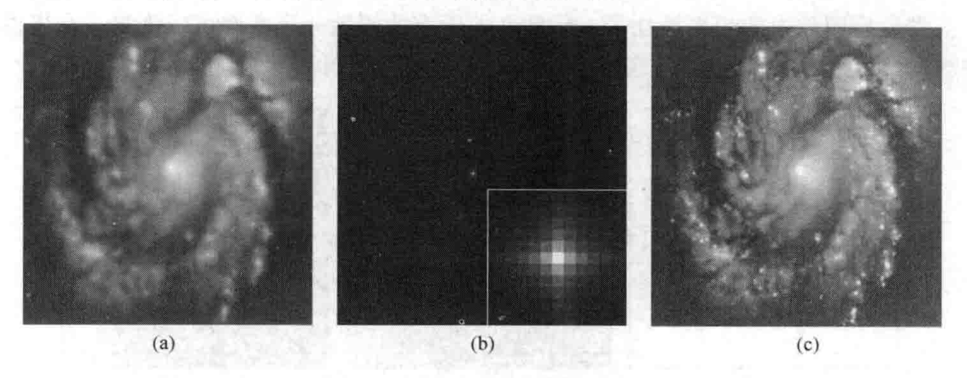


图 6.37 哈勃太空望远镜图像锐化：(a)原始图像；(b)所测点扩散函数，放大部分显示了像素细节；(c)去卷积结果

下面对灰度图像讨论去卷积。多数情形下，彩色图像需要分离为对应于获取设备的各个彩色通道，通常为红色、绿色和蓝色通道，每个通道再单独去卷积。通常，求每个通道的点扩散函数的方法是不同的，尤其是获取设备是带有拜尔彩色滤镜的单芯片相机时。

总是从尽可能好的图像开始很重要。这意味着图像获取时要有最好的聚焦、最少的运动模糊等。去卷积从来不是校正问题来源的好方法。图像记录时，应有最好的对比度范围，即从近乎全白到近乎全黑。色调或灰度范围应有较好的精度和较宽的动态范围。非冷却式数码相机获得的 8 位图像基本合适，摄像机获得的图像通常不可接受，除非使用专用程序进行了处理（如平均多个帧）。尤其重要的是，无论是通过测量还是通过计算得到的，点扩散函数都要有较高的精度和位深。

最后，随机像素噪声（斑点）应被降到最低。无论是已获取图像中的噪声，还是点扩散函数中的噪声，在去卷积过程中都会明显放大，且如果噪声太大的话，那么会在结果中占支配地位。长时间曝光和图像平均操作在某些情形下是有用的。图像噪声的振幅通常描述为以分贝表示的信噪比。这是根据模糊图像和噪声中值的标准差（这些值可能并不知道）来定义的：

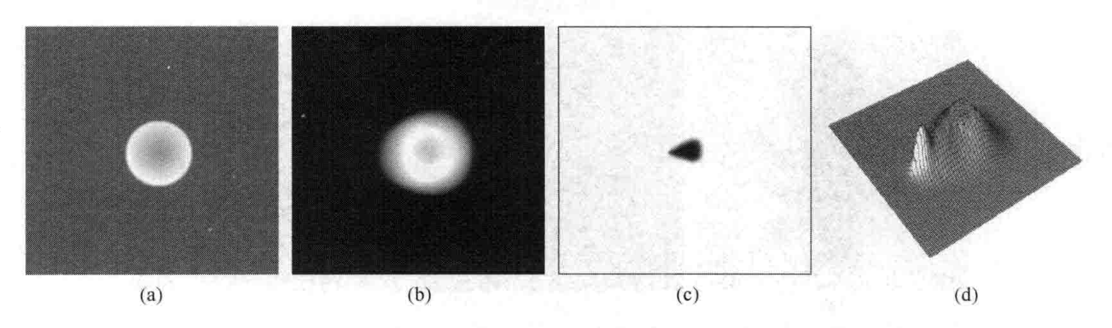
$$\text{SNR}\,[\text{dB}] = 10 \cdot \lg\left(\frac{\sigma_{\text{图像}}}{\sigma_{\text{噪声}}}\right) \tag{6.11}$$

当信噪比大于 50dB 时，噪声在图像中实际上并不明显，且对去卷积的影响很少。另一方面，10～20dB 的低信噪比会使得噪声非常显著，以致不适用于去卷积。

去卷积的理想和最简形式是测量（或在少数情形下由已知的光学参数进行计算）系统的点扩散函数。首先计算模糊图像的傅里叶变换和 PSF 的傅里叶变换，然后用后者除以前者，再进行反变换，就会生成去卷积的结果。关键要求是，假定成像系统所致的模糊在图像中的各处都相同，对望远镜图像来说这是一个较好的假设。当部分模糊由光通过样本本身所致时（如在光学显微镜下检查厚切片时出现的情形），该模糊会逐点变化，此时就不能使用这种方法了。

使用与真实图像所用的相同光学设置来获取完全已知物体的一幅图像，可间接测量 PSF。将物体图像的傅里叶变换除以理想形状的傅里叶变换，然后对结果反变换，可得到系统的 PSF，进而使用它来去卷积由光学系统获得的图像。

在有些显微镜情形下，可通过在幻灯片上插入荧光磁珠或一幅尘埃图像来计算 PSF。在原子力显微镜中，通过扫描一幅形状（通常使用一个圆形测试图案，其生成方法与刻蚀集成电路的方法相同）已知的图像，可直接测量 PSF。如图 6.38 所示，如果所得图像的傅里叶变换除以理想形状的傅里叶变换，那么结果是点扩散函数的一幅图像，此时该图像对应于扫描探针的形状。然后可用这种形状来对其他图像去卷积，直到该探针被损坏或被替换。



(a) (b) (c) (d)

图 6.38　测量原子力显微镜的 PSF：(a)圆形测试图案的理想图像；(b)使用有缺陷探针测量的图像；(c)图(b)被图(a)反卷积，生成了探针的图像；(d)图(c)形状的透视图

纯光学效应所致模糊在各个方向经常是一致的，尽管像散可能会改变这一点。在许多相机中，传感器形状和间距在水平和垂直方向上会产生不同的模糊振幅，且所用滤色镜模式的影响也会改变 PSF 的形状。获取任何扫描图像时，所用放大器的电子参数，会在快速扫描方向（通常为水平方向）和慢速扫描方向（垂直方向）上产生不同的模糊量。荧光粉、放大器或其他组件的时间常数在输出信号中也可生成不对称模糊（“彗尾”）。

如果通过测量不能得到 PSF，那么可基于假设（如高斯模糊函数、表示相机光圈的圆盘或直线运动）来应用一个近似的 PSF。在许多情形下，无须精确的 PSF，大致近似的 PSF 就可通过去卷积操作来消除大部分模糊。

6.12　噪声和维纳去卷积

如果待锐化图像中存在明显的噪声，或所测 PSF 中存在明显的噪声，则可能会加剧数值精度和溢出问题，并使反变换结果严重退化（Vanderlinde & Caron，2008）。删除真实图像中的大部分模糊几乎不可能，但在有些取证情形下，即使是很小的改进，也会有相当大的实用价值。

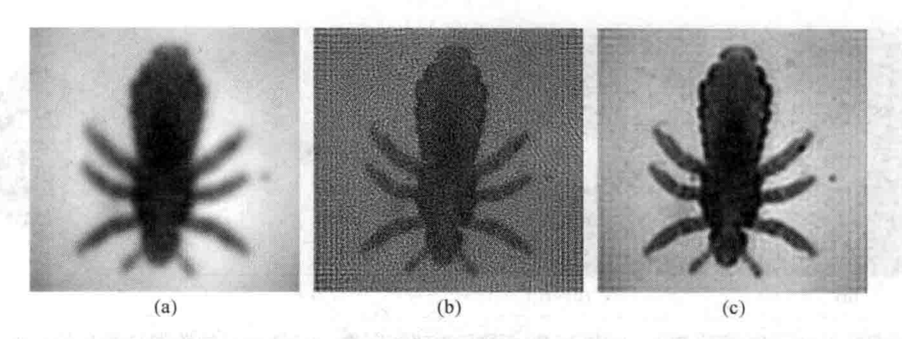
除以模糊的频率变换称为逆滤波器。记图像为 $f(x, y)$、模糊图像为 $g(x, y)$、点扩散函数为 $h(x, y)$，它们的傅里叶变换分别是 $F(u, v)$、$G(u, v)$和 $H(u, v)$，则“理想”去卷积过程可以写成

$$F(u,v) \approx \left[\frac{1}{H(u,v)}\right] G(u,v) \tag{6.12}$$

如果除以傅里叶变换的去卷积运算因模糊图像中的噪声而不能得到令人满意的结果，则实用方法是进行维纳去卷积运算。前者使用原始图像的傅里叶变换除以模糊函数的傅里叶变换来计算去模糊图像，而后者使用一个标量值来增大分母。从理论上讲，加性因子 K 取决于图像的统计特性及它们的相对噪声内容（它是图像噪声分量功率与图像信号分量功率之比），但实际上这些通常都是未知的，因此加性常数通常被视为一个在锐化和噪声间保持平衡的可调参数。

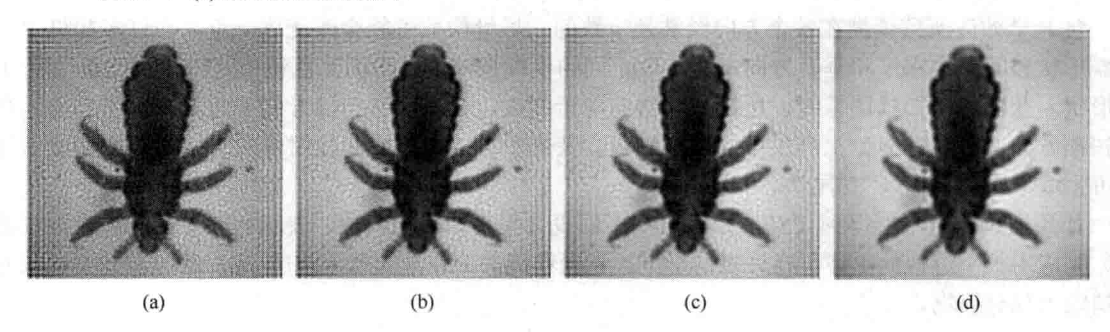
$$F(u,v) \approx \left[\frac{1}{H(u,v)}\right] \cdot \left[\frac{|H(u,v)|^2}{|H(u,v)|^2 + K}\right] \cdot G(u,v) \tag{6.13}$$

图 6.39 显示了随机噪声对一幅因散焦而模糊的图像的影响。使用一个理想逆滤波器放大噪声，以便它在结果中占主导地位，并掩盖任何有用的细节，虽然边缘变得清晰了。切趾法（跳过那些除法将产生数值溢出的频率项）产生了更好的结果，但出现了振铃现象，且残留了一些模糊。维纳去卷积（见图 6.40）的结果更好，并允许对结果进行调整，即在清晰度的恢复（高 K 值留下较多的模糊）和噪声量（高 K 值降低噪声）之间保持平衡。由于原始图像的信噪比较低，且存在大量的模糊，因此此时不可能恢复原始图像中的所有内部细节。



(a) (b) (c)

图 6.39　噪声对去卷积的影响：(a)严重模糊的图像［加有噪声的图 6.34(c)］；(b)理想的反去卷积结果，存在噪声；(c)切趾法反卷积结果



(a) (b) (c) (d)

图 6.40　图 6.39(a)所示图像的维纳去卷积：逐渐增大 K 值，会在牺牲清晰度的代价下，降低噪声

作为一种消除模糊的实用方法（尽管并不完美），维纳去卷积已有很长的历史。当信号（图像）被随机加性噪声模糊或破坏时，维纳方法是最佳的恢复技术，因为这种方法会在统计意义上最大限度地减少实际（真实）图像与重建结果间的误差。另一种知名方法是 Van Cittert 迭代技术。这种方法显然已被多次发现，因此有时也称为 Bially 或 Landweber 迭代。这种方法不是一步来消除整个模糊，而是要执行一系列的迭代消除步骤。这种迭代的结果与理想去卷积的结果相同，但它会在收敛之前终止，并显示一幅噪声水平可接受的去模糊图像。所得结果的质量通常类似于维纳方法结果的质量。

去卷积的方法有许多，且多数方法的数学计算相当复杂（且需要大量的计算和时间）。有些是非迭代的（如 Carasso，2001，2002，2006；Zhang & Zhang，2007），但多数会涉及迭代，如 Lucy-Richardson（Richardson，1972；Lucy，1974）、Landweber（Landweber，1951；Vonesch & Unser，2008）和 Van Cittert（Van Cittert，1931；Bennia & Riad，1992）方法。有些方法假定 PSF 是已知的，有些方法则假定 PSF 是对称的，有些方法会试图找到最佳形式的模糊函数和去模糊图像。

这通常等效于试图求解一个非常大的欠定联立方程组。在许多数学文献中可找到此类方案的技术，在第 13 章关于三维成像的内容中，也会介绍这种技术。问题之一是如何有效地引导迭代及何时终止它。另一个相关的问题是图像质量的最佳测量，在不同的实现中，它可能会基于贝叶斯统计、熵的最大化等。计算可在频率域中进行，也可在空间域中进行。即使空间域中的一个小 PSF，在频率域中也会变得与整幅图像一样大，且联立方程的个数是图像中像素的数目。

多数此类方法是针对某类特殊应用的，且依赖于可以提供的先验信息，即方程组的约束条件（Bertero & Boccacci，1998）。例如，常用的约束条件是负值像素无物理意义，因此值不能是负的。如果可得到关于图像的更多信息，如大部分图像区域的背景恒定（就像在天文学或荧光图像中那样），或仅允许对应于相位密度的特定像素值，就会提高结果的质量和实现这种结果的效率。这些方法的详细内容超出了本书的范围，读者可在 Lagendijk & Biemond（1991）和 Campisi & Eglaarian（2007）中找到关于它们的优秀回顾和比较。

当前关心的另一个领域是彼此关联的多幅图像的去卷积。彼此关系图像包括同一场景的多个通道图像（如不同的波长或颜色），或半透明标本的一系列排列紧密的图像（如光学显微镜中的一系列焦平面图像）。一个通道或平面的信息可用来指导另一个通道或平面的去卷积。

如果事先并不知道模糊，那么可由功率谱估算，或由试错法找到某个最佳（或至少有用）的结果。维纳去卷积的速度很快，因此可交互式地调整常数 K 和 PSF 函数的形状。尖峰较尖的函数可为真实图像产生更好的结果。图 6.41 显示了一个例子，并与高斯函数的结果进行了比较。

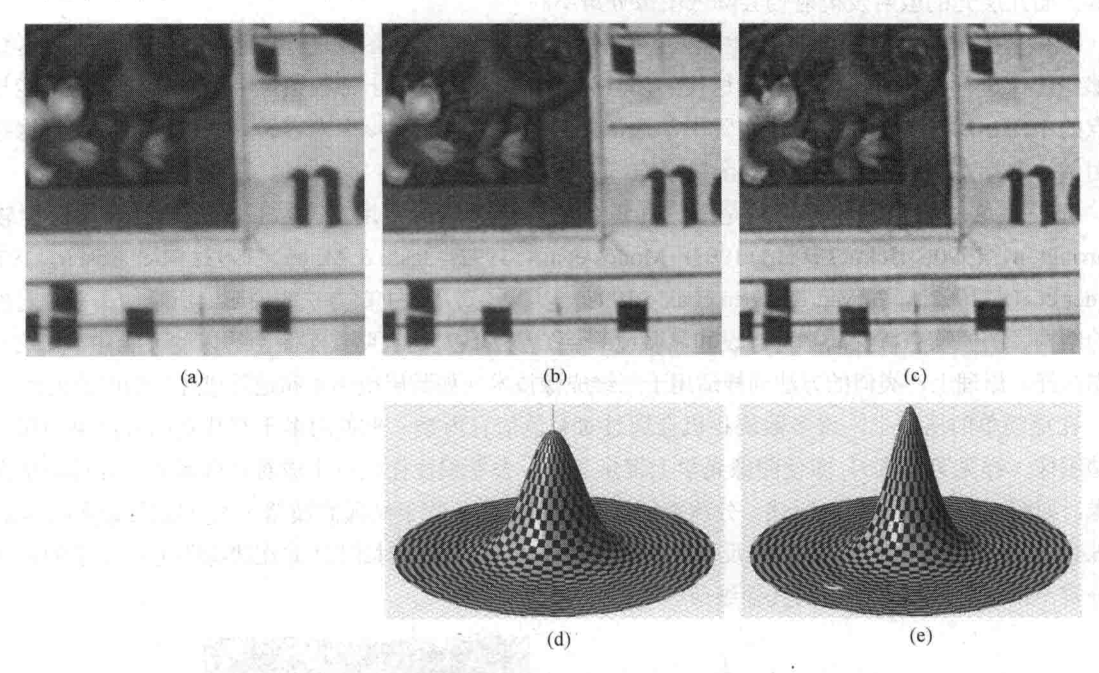


图 6.41 使用不同 PSF 函数去卷积结果的比较：(a)原始图像（带有明显像素噪声的手稿）；(b)使用图(d)中所示高斯 PSF 的最好结果；(c)使用图(e)中所示函数的最好结果；(d)高斯 PSF 形状 $z=\exp(-r^2)$；(e)更尖尖峰的数学函数 $z=\exp(-r^{1.7})$

此类函数的一种最新发展已具有很高的效率，因为它不是由 Carasso（2001，2002，2006）发布并应用于哈勃望远镜图像的迭代方法。分析模糊图像（避开噪声非常多的区域）的傅里叶变换中值的径向剖面形状，可构造一个可用于去卷积的近似 PSF。这种方法不能用于所有类型的模糊函数（尤其是运动模糊），但对那些适用情形能产生非常好的结果。

与图 6.41 中所示函数相比，稍微模糊的盘状（可以是圆形或多边形，以匹配相机的光圈）PSF，对应于某种情形的散焦镜头模糊。图 6.42 显示了一个例子，其中 PSF 是一个稍微模糊的六边形。

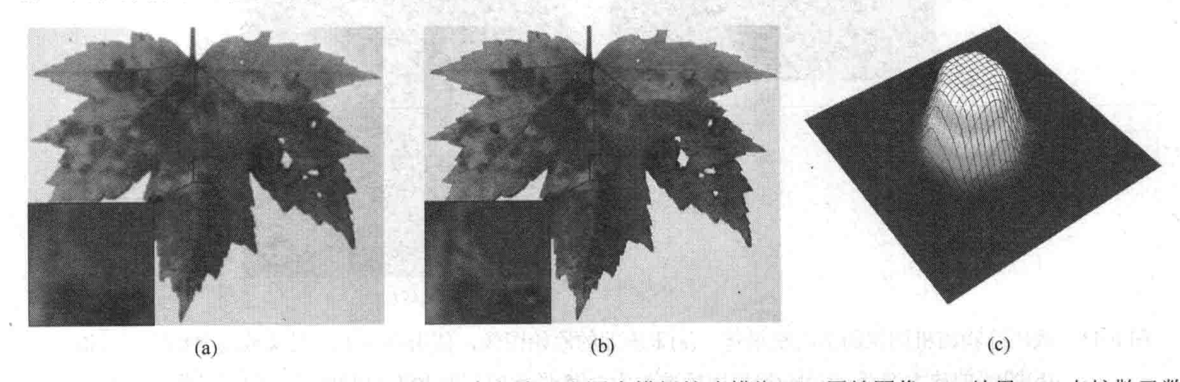


图 6.42 使用一个 PSF 去卷积，该 PSF 用一个六边形光圈来模拟镜头模糊：(a)原始图像；(b)结果；(c)点扩散函数

这些图像恢复技术的一种应用是，去除光学切片图像中的模糊。光学切片成像是使用光学显微镜，对半透明标本采用不同焦深来记录一系列图像的技术。穿透标本底层的光会导致某种模糊，这种模糊会影响到图像的清晰度和对比度，因此不能将它们放到三维堆叠图，进而测量三维结构图。

共焦光学显微镜克服了某些此类问题，方法是拒绝来自偏离焦点的那些点的光，这增强了图像的对比度（但也降低了光学景深，在深度轴上产生更高的分辨率，在最高放大率情形下这相当重要）。但标本上面几层光的散射及衍射仍会降低图像分辨率。

从原理上讲，上面几层的图像中包含有可用于对下面几层的图像进行去卷积操作的信息。这允许对那些图像进行锐化。整个过程是迭代的，且计算开销较大，由于每幅图像上每个点的模糊可能与其他点的不同，情况更是如此。在实际中，通常会对图像的模糊和噪声内容做一些假设，然后将这些假设用做给定标本或给定光学设置的全局平均。

即使有这些假设，计算仍然是密集和迭代的。该领域的实用文献数量不多，且理论深度较大（Carrington，1990；Holmes et al.，1991；Monck et al.，1992；Joshi & Miller，1993；Richardson，1972；Snyder et al.，1992）。在 Van Kempen et al.（1997）中可找到几种主要方法的说明。所示例子只针对理想化的结构、图像噪声特性的平均假设和显微镜的点扩散函数，这说明真实图像的恢复可能并不如这些例子那么好。原理上，类似的方法同样适用于三维成像技术（如断层摄影术和地震成像）的图像处理。

在更简单的层面上，许多数码相机会通过对彩色信息内插，来使用多于芯片实际分辨率的像素来存储图像（详见第 1 章）。即使图像光学上聚焦良好，去卷积往往也可生成对比度增强、分辨率增大的图像，如图 6.43 所示。如前所述，分辨率的这种提高完全不同于高通滤波器（如反锐化掩模）的结果。图 6.44 比较了这两个过程的结果。反锐化掩模增强了原有边缘的对比度（并在边缘附近产生了的晕环），但分辨不出去卷积操作产生的其他细节。

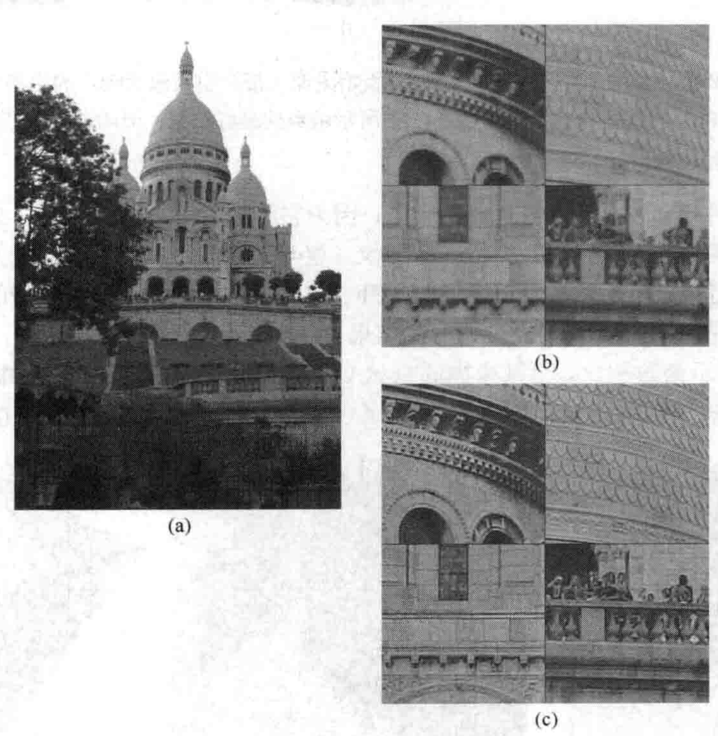


图 6.43　数码静物相机图像的去卷积锐化：(a)未压缩的原始图像，使用高端单反尼康数码静物相机拍摄；(b)图(a)的放大细节；(c)图(b)使用标准差为 0.95 像素的高斯 PSF 进行维纳去卷积操作后的结果

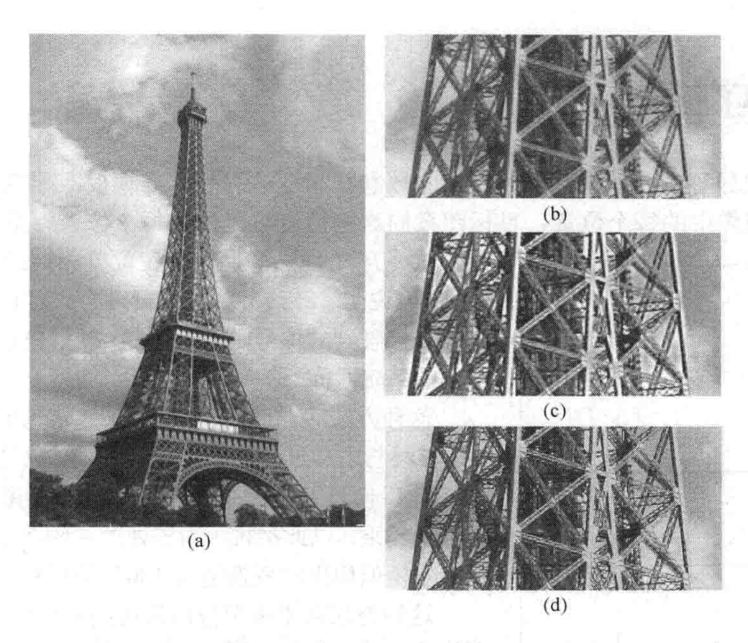


图 6.44　应用高通滤波器的去卷积比较：(a)原始图像；(b)图(a)的放大细节；(c)图(b)使用反锐化掩模后的结果；(d)图(b)经维纳去卷积后的结果

去卷积运算除了可校正散焦光学问题之外，还可校正其他缺陷。这些运算并不总是在频率域中执行，但清楚说明了系统强加卷积操作来消除模糊的过程。最常见的一种缺陷是由运动引起的模糊。在遥感成像（此时光线不充足，且在相机运动相对于场景的运动很明显时，需要更长的曝光时间）中，运动引起的模糊很明显。

所幸的是，在多数此类情形下，运动的距离和方向是已知的，因此可在空间域中画出一条定义模糊的直线。如果画不出这样一条直线，那么可使用一个亮特征或暗特征来测量出模糊的角度和方向，进而画出这条直线。然后用这条直线的频率变换除以模糊图像的变换。反变换结果图像，可恢复出清晰的结果。图 6.45 示例了这种可能性。注意，在该例中，相机曝光期间的轻微旋转运动，造成了图像的顶部和底部边缘未完全恢复。该去卷积过程假设对整个图像区域应用了相同的 PSF；旋转的去卷积需要一种单独的方法（Yuan & Li，2008）。

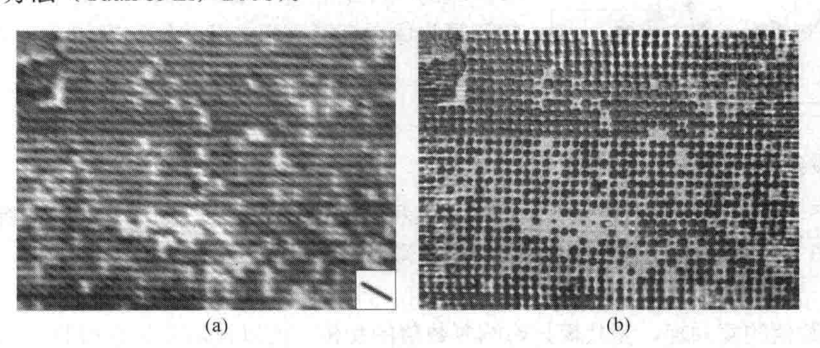
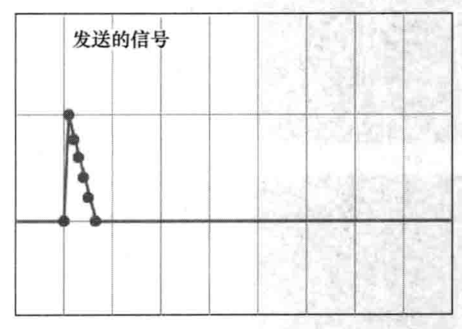


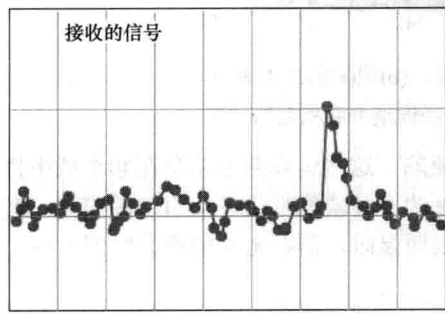
图 6.45　运动模糊的消除：(a)原始图像（果园的航拍照片），其上叠加有（放大后）的运动矢量；(b)维纳去卷积后的结果

请特别注意该例与散焦模糊消除间的相似性和差异。二者均涉及用模糊图像的变换除以缺陷的变换。这与直接由卷积方程得到的结果相同，在卷积方程方法中，卷积图像的变换是原始图像变换和缺陷变换的积。主要区别是，在运动模糊情形下，有时可能知道精确的待消除模糊矢量，或至少可由图像估计出模糊矢量。如上面关于散焦模糊的例子所示，使用维纳去卷积可部分减轻噪声的影响。

6.13 模板匹配和相关

与为平滑、求导等应用一个核来进行空间域卷积相近的思想，是模板匹配或互相关。此时，目标图案会移动到图像中的每个位置，目标图案的这些值会与覆盖的每个像素相乘，乘积之和存储在该位置，形成一幅与目标所在位置相同或类似区域的图像。图 6.46 说明了该过程的一维处理方式：由雷达发射器发出的脉冲与接收到的信号是互相关的，以指出该信号传播到目标并返回的精确时间。





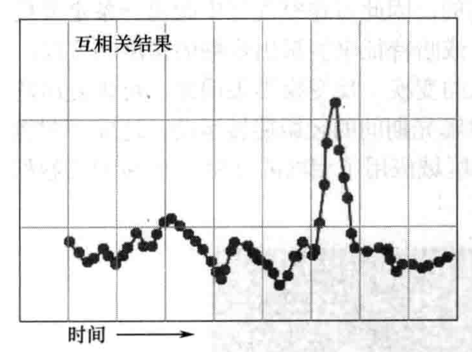


图 6.46 雷达发射机中使用的一维互相关。发送出的脉冲波形与接收到的信号互相关，以求出时间延迟，进而求出到目标的距离

这种方法同样适用于图像，只是它基于两个空间度而非时间。相乘与求和过程与卷积是相同的，只是首先要将目标旋转 180°，以便目标图像中的左上角点值与图像上邻域中的右下角点值相乘，以此类推。当该处理在频率空间中执行时，与卷积等效，但傅里叶变换值有 180° 的相移。

这种方法常用来定位图像内的各个特征，如搜索侦察图像中的特定目标（如车辆），在一系列气象卫星图像中跟踪飓风的运动，在显微镜幻灯片上跟踪细胞的移动等。这种方法针对二值图像的一种最优改进，可用于查找文本中的字母。当目标是立体像对中一幅图像的像素亮度值图案且所搜索图像是该立体像对的第二幅图像时，可使用该方法来执行融合操作（即在两幅图像中定位匹配点），进而测量视差并计算高程，如第 14 章中的图 14.9 所示。

对于连续的二维函数，互相关图像计算为

$$c(i,j)=\iint f(x,y)\cdot g(x-i,y-j)\mathrm{d}x\mathrm{d}y \tag{6.14}$$

在图像的各个维度上，通过用有限求来代替积分，可得式 (6.15)。为归一化该模板匹配或相关的结果，且在确保图像中该区域的亮度绝对值不偏离结果的前提下，空间域中该运算通常计算为像素亮度乘积之和除以它们的几何平均：

$$\frac{\sum_{i,j} f_{x+i,y+j}\cdot g_{i,j}}{\sqrt{\sum_{i,j} f_{x+i,y+j}^2\cdot\sum_{i,j} g_{i,j}^2}} \tag{6.15}$$

当求和的规模较大时，与频率空间中的等效运算相比，这是一个缓慢而低效的过程。频率空间运算是

$$C(u,v)=F(u,v)G^*(u,v) \tag{6.16}$$

式中，*表示函数值的复共轭。复共轭只影响复数值的相位，所以该运算与卷积非常相似，且它通常由许多相同的子程序来执行。涉及两幅图像的运算（用于去卷积的除法、用于卷积的相乘和用于相关的共轭相乘）有时也称为二元运算，以将它们与只针对单幅图像执行的滤波和掩模运算区分开。滤波器或掩模图像不是频率域图像，不包含复数值，并且只与图像变换的振幅相乘。

执行相关运算时，连接图像左右边缘和上下边缘的环绕假设通常是不可接受的。此时，应将每幅图像填充到适合于变换的更大尺寸（许多 FFT 程序所需要的尺寸是 2 的整数次幂）。由于相关运算要求使用变换的实际振幅值，因此使用原始图像的平均亮度值而非零值来进行填充，数学精度会更

高。从每个像素的亮度值中减去平均亮度值也是有用的，这会从变换中消除第零（直流）项。由于该值通常是变换中的最大值（即中心像素的值），消除它可使得变换数据有更大的动态范围。

相关运算主要用于查找两幅图像中均出现的特征（Caves et al.，1992；Yu et al.，2004）。图 6.47 给出了一个例子。该图像中包含带有一些随机噪声的文本，而目标本身包含了字母“A”。互相关的结果（将图像反变换到空间域后）显示了对应于目标字母的各个尖峰，该图像等距显示时，各个尖峰更为明显。相关图像中的最暗点对应于大小和字体与目标相同的字母“A”。较小但仍非常明显的尖峰对应于其他两个不同字体的字母 A，但互相关通常适用于具体的大小和形状。

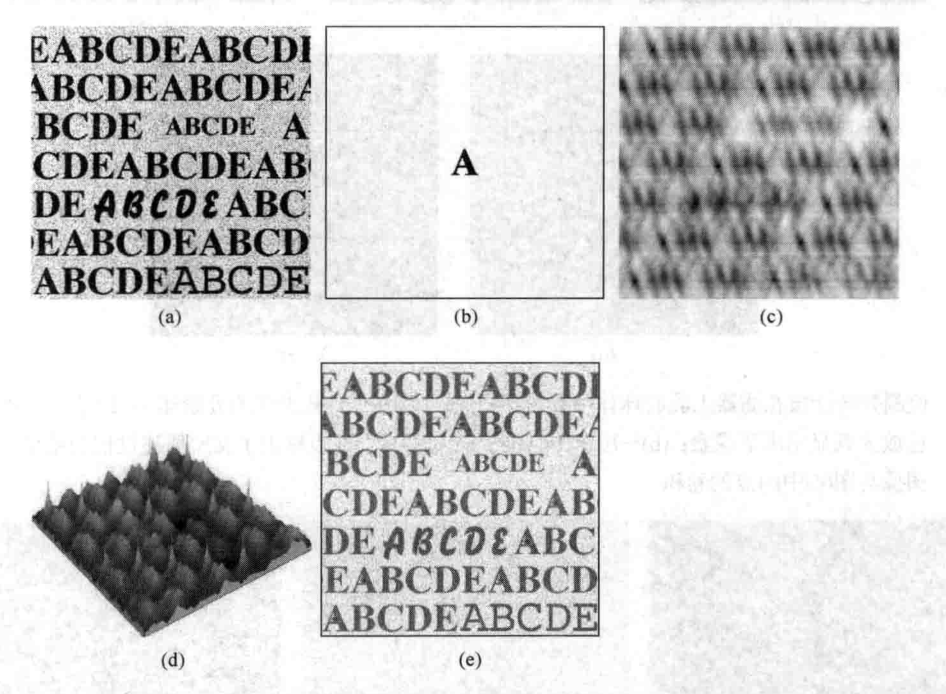


图 6.47　互相关示例：(a)包含许多字母且叠置有随机噪声的图像；(b)目标字母；(c)互相关结果（良好匹配的灰度表示）；(d)图(c)的等距视图，它显示了原始图像中字母 A 上的尖峰；(e)图(c)经阈值处理后叠加到图(a)上的结果。注意，仅找到了相同大小和字体的字母 A

结合使用其他图像处理工具来分析互相关图像时，这是定位图像中已知特征的一种快速且有效的工具。在图 6.48 所示例子中，由于对比度和噪声的变化、灰尘的存在以及滤波器本身的纹理，导致在滤波器的 SEM 图像中自动计算乳胶球体的数量很是困难。此外，一个孤立球体的对比度与被其他球体包围的那个球体的对比度是不同的。

平均 10 个有代表性的乳胶球体，创建了一幅目标图像［见图 6.48(a)］。操作步骤如下：与每幅图像互相关，对图像中的每个尖峰应用一个顶帽滤波器（内径为 3 像素、外径为 5 像素、高度为 8 个灰度值），自动计算出点数。该图显示了图像集中的几幅例子图像，并使用标记显示出了所找到的颗粒并统计了颗粒的个数。应用顶帽滤波器选取了互相关图像中的各个尖峰后，得到的点与存储的目标图像卷积，生成如图 6.48(e)所示的结果，该图像只显示了原始图像中的乳胶球体而未显示其他特征（如灰尘、背景纹理等）。

某些情形下，可在图像中找到目标例子并用于搜索类似的内容。图 6.49 显示了涂漆面上的气泡。光照来自一侧且图像是带噪的，因此气泡有复杂的阴影。用一个（标有红色边框）气泡作为目标并对整幅图像执行互相关运算，然后应用一个顶帽滤波器，可找到其他气泡，即使它们具有不完全相同的阴影和大小。

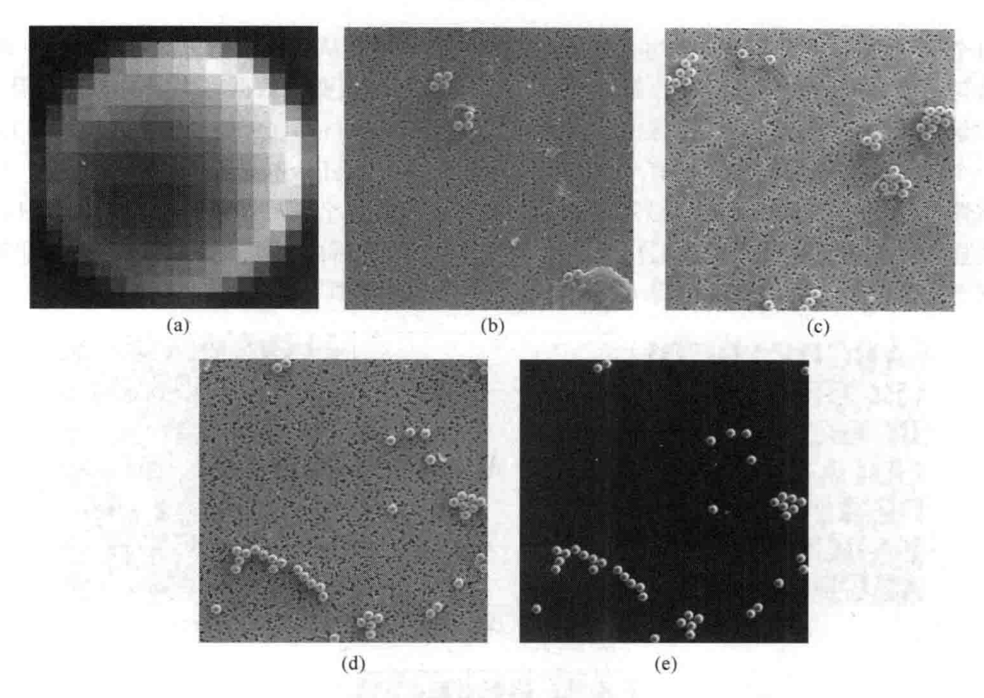


图 6.48 检测并统计核孔滤器上乳胶球体个数的互相关：(a)10 个有代表性的乳胶球体的平均，用做目标，已放大到显示出了像素；(b)～(d)有代表性的示例图像，红点标出了文中所述过程的结果；(e)目标图像与图(d)中的点的卷积

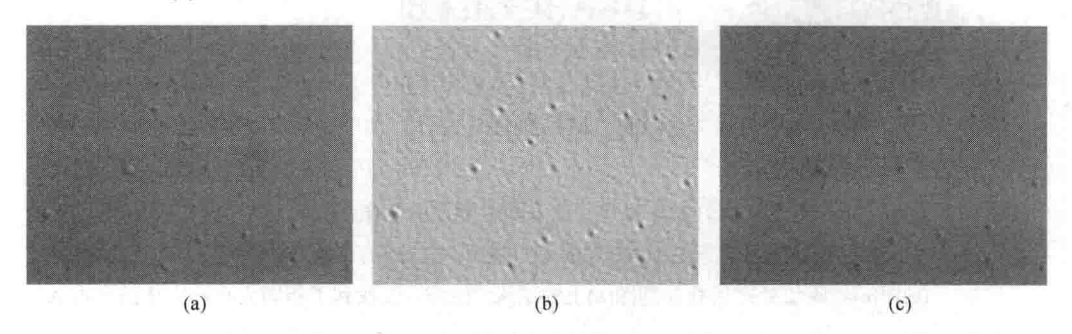


图 6.49 检测油漆中气泡的互相关：(a)原始图像，圈出了用于互相关的目标例子；在气泡上显示了暗色尖峰的互相关结果；(c)顶帽滤波器叠加到原始图像上的结果

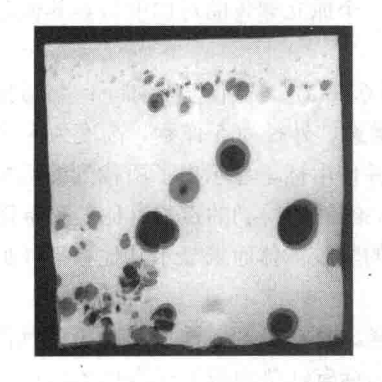


图 6.50 瑞士奶酪的两幅连续切片图像通过互相关自动对齐并叠加后的图像

互相关也被用做对齐图像的一种工具（Reddy et al.，1996；Choi & Kim，2002）。即使图像不相同（例如对于连续切片），但当两幅连续图像在 x 和 y 方向较为对齐时，通常也会出现足够多的特征来锐化互相关尖峰。尖峰的位置可精确到子像素精度，并用于移动图像来得到最好的对齐。图 6.50 显示了一个例子。瑞士奶酪的两个连续切片中，各个小孔的位置和大小有一些细微的差别，而在一个切片中出现的几个小孔，在另一个切片中则不存在，这对第 9 章中介绍的体视测量工具（Disector）非常重要。但在两个切片如图所示那样完全对齐前，比较是难以进行的。

对于第 14 章中介绍的所有重建方法，对齐都是重要的。如第 5 章所示，将多幅焦平面图像合并为一幅扩展的聚焦图

像，对齐也是必要的。图 6.51 是来自光学显微镜的一系列焦平面图像，各连续平面图像间都存在偏移，原因在于显微镜的机械和光学影响。在图 6.51(c)的红色和绿色通道中，叠加经边缘滤波的图像（使用第 5 章的 Sobel 边缘描绘滤波器）后，结果显示了这一偏移。互相关运算可将每个平面的偏移自动校正到子像素精度，因此可由这些图像形成一幅扩充的聚焦图像。在应用 Sobel 滤波器之后，对图像应用互相关操作产生了更好的结果，因为成功对齐所需的多数信息都位于边缘中。

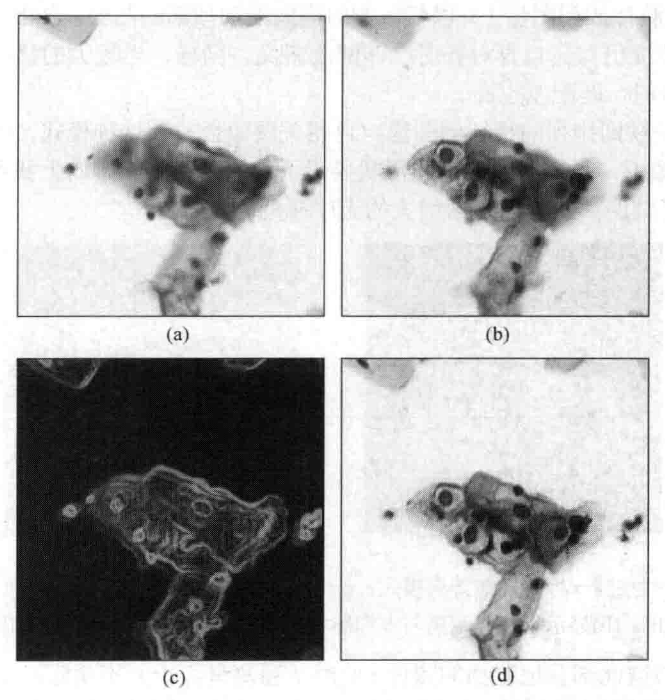


图 6.51　一系列焦平面图像的对齐：(a)和(b)具有不同景深的原始光学显微镜图像（顶部和底部聚焦）；(c)红色和绿色通道中（经边缘滤波后）的两幅图像的叠加，显示了横向偏移；(d)所有图像对齐后的扩展聚焦图像

最严格的对齐任务之一是遥感图像的对齐。遥感图像通常是由不同相机和其他设备在不同时间获取的，伴随着比例尺和方向的变化。对齐通常是基于每幅图像中的几个参考点的位置而非校正来完成的。该对齐过程可能涉及图像扭曲来校正不同的视角。

标准视频由含有奇数和偶数扫描线的交变场组成。如果相机快速摇动镜头或图像中的物体正在移动（监控录像通常属于这种情形），就会在各场间产生偏移，且这种偏移会显示在数字化图像中，如图 6.52 所示。通过将这两个交变场（即偶数和奇数扫描线）当做两幅单独的图像来处理，互相关可用于校正这一偏移，如图 6.52(b)所示。

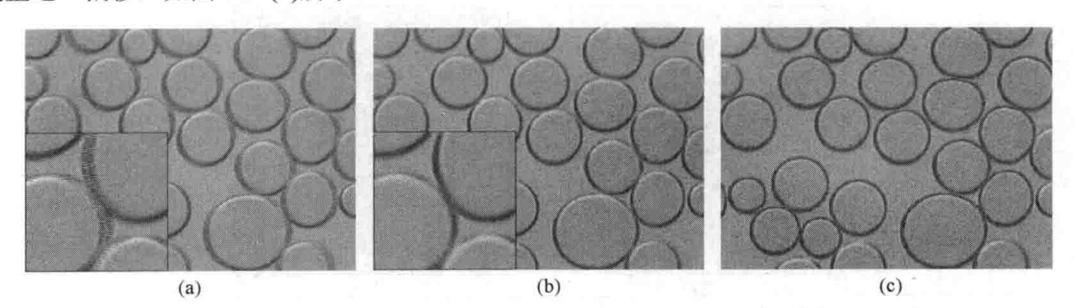


图 6.52　对视频图像去隔行扫描：(a)气泡在流动通道中移动的原始图像，产生了偶扫描线和奇扫描线偏移（插图显示了放大的细节）；(b)校正后，各交变场自动对齐后的结果；(c)（使用第 7 章和第 8 章中的步骤）对二值图像阈值处理后，形成了便于测量的气泡轮廓

6.14 自相关

在图像函数 f 和 g（及其变换 F 和 G）相同的特殊情况下，相关运算称为自相关。它用于将图像的所有部分合并到一起，进而找到重复的结构。自相关图像的理解如下：将图像印到透明胶片上，然后将旋转 180° 后的透明胶片放到图像上。以任何方向横向滑动顶部的图像，自相关函数就可测出其与底部图像的匹配度。当特征仍与其自身对齐时，匹配度就高。同样，当较大的移动使得某个特征出现在另一个类似的特征上方时，匹配度也高。

对于其中存在一些规则排列的特征的图像，自相关图像会找到这种模式，因此也会便于测量，如图 6.53 所示。在该图像中，每个点的暗度和形状显示了在给定距离和方向上找到另一个颗粒的概率。圆形平均径向图显示了随距离变化而逐渐增大的无序排列。

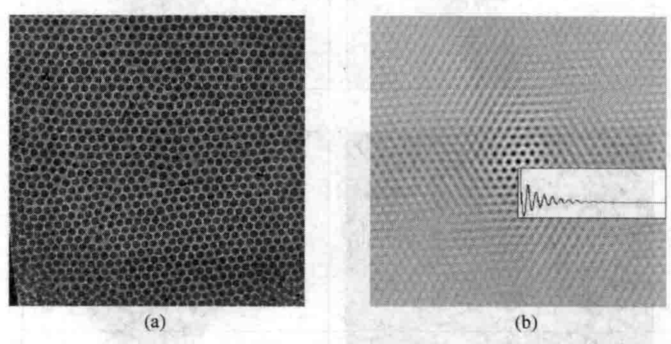


图 6.53　使用自相关找到一种有规律排列模式：(a)乳胶球体的 TEM 图像，它是以近似六边形图案来自我组织的；(b)显示了相邻距离与方向图案的自相关图像，插图是圆形平均径向曲线的度量

对于其中各个微粒彼此部分地遮挡的图像（此时单独测量大小并不实际），自相关有时可提供一种简单但强大的方式来测量大小、形状甚至粒度分布。图 6.54 显示了使用不同 pH 值制作的两种不同奶

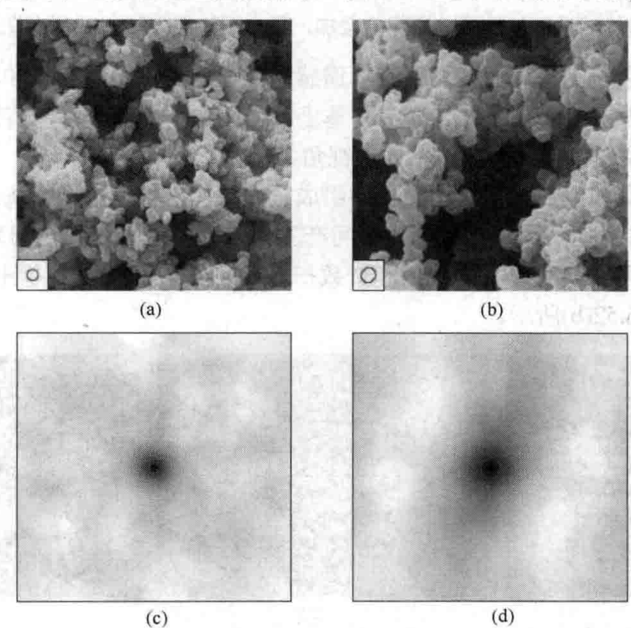


图 6.54　重叠特征的测量：(a)和(b)以不同 pH 值和不同凝乳尺寸制作的两个奶酪的 SEM 图像（插图显示了对图(c)和(d)中各尖峰进行阈值处理后，获得的平均特征尺寸的轮廓）；(c)和(d)图(a)和(b)的自相关结果，等值线画在最大值的 50%处

酪的表面。每幅图像中都显示了大量且几乎相同的元素。视觉上，两个标本中的凝乳大小存在差异，但由于这是一个不规则的表面，且微粒是堆积的而不是分散的，造成许多微粒至少部分地被其他微粒遮挡，因此进行传统的特征测量是不可能。另外，还显示出了自相关图像。

中央尖峰表示特征不再位于其本身上方前，图像可以横向移动的距离，因此使用中央尖峰的大小就可方便地测量奶酪中结构元素的大小。图中，通过在自相关图像上的50%振幅处画一条等值线，显示了两个结构来表示平均尺寸。注意，自相关图像是实空间图像，因此该尺寸的含义与原始图像中的相同。尖峰的亮度剖面还包含有关于微粒大小分布的信息。

在第 15 章中，自相关将根据尖峰下降幅度的特征距离，来测量表面粗糙度的纹理。图 6.55 显示了自相关函数的另一种用途，即测量首选方向。该图像是一种毡制纺织材料。由于制造工艺的原因，纤维没有随机的方向或统一的方向；它们也不是有规律地排列的。沿其长轴和短轴的自相关中心点的亮度剖面，定量地测量了首选方向。自相关函数的径向剖面已被解读来估计砂岩渗透率（Berryman and Blair，1986）。

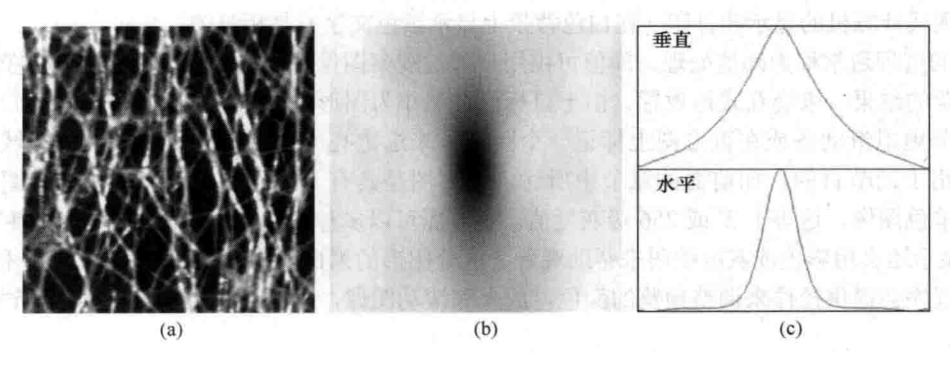


图 6.55 首选方向的测量：(a)毡制纺织纤维图像；(b)自相关尖峰（已放大）；(c)尖峰的水平和垂直亮度剖面

第 7 章　分割和阈值处理

7.1　阈值处理

在一个场景或一幅图像中选择物体或特征，是进行多数类型测量或分析的重要前提。传统上，完成这种选择的一种简单方法是，在原始图像中定义一个亮度值范围，只要像素落在这个范围，则将其划归为前景，而将其他像素划归为背景。然后将这样一幅图像显示为二值或双电平图像，使用黑色和白色（有时也用其他颜色）来区分不同的区域。感兴趣特征是白色还是黑色，并不存在标准约定；所做的选择取决于所用的特殊显示硬件和个人偏好。在此处显示的例子中，特征为黑色而背景为白色，这与多数现代计算机的显示和打印（在白色背景上显示黑色文字）是相符的。

选择的过程通常称为阈值处理。阈值可由用户通过观察图像并使用彩色叠加来显示调整按钮或调整其他设置的结果，来交互式地设置。由于鼠标已普遍作为图形化计算机显示器的人机接口，因此用户可以调节虚拟滑动条或在直方图上标记一个区域，来选定亮度值的范围。图像（或区域）的亮度直方图常用于调节目的。如前面的章节中所述，直方图是具有不同亮度级图像中的像素图。对于典型的 8 位单色图像，这等于 2^8 或 256 级灰度值。直方图可以多种格式呈现，如垂直直方图和水平直方图，有些显示还会用彩色或灰度编码来帮助观看者区分图形的黑白两侧。图 7.1 说明了该技术的基本原理；很多程序会提供控件来调整预览的颜色，放大和滚动图像，有时还可自动设置阈值，详见本章后面的内容。

具有 256 级（8 位每颜色通道）以上色调范围的成像系统（如扫描仪、某些数码相机和其他设备）所生成的图像，可保存为 16 位文件（65536 个不同的像素值），尽管多数情形下这并不重要。就方便性与一致性而言，都是为了阈值处理目的，以及为了保留前面章节中所介绍的各种数值常量的含义，这样的图像仍然可以描述为具有 0～255 亮度值范围的图像，以便覆盖从黑到白的范围。

在不限于整数亮度值的情形下，更精确的数据允许亮度为实数（例如，65536 个值中的像素值 31605 除以 256，得到 123.457）。这便于那些与原始图像深度和动态范围无关的值的比较，但未解决如何在直方图中显示这样一个范围的值的问题。包含超过 65000 个值的完整直方图对于计算机屏幕而言太宽，而且对于典型的图像大小而言，每个通道中非常少的计数会导致人们难以解读。一般来说，为便于查看，这些计数会分段到更小数量的通道上，但对于 8 位图像仍然多于 256 个。图 7.2 显示了额外的细节，出现更多的尖峰表明可为直方图使用更多的通道。

前几章中给出了直方图的几个例子。直方图统计整幅图像（或感兴趣区域）中的像素数量，而与像素的原始位置或它们的相邻像素的亮度值无关。直方图中的尖峰识别不同的均质区域（通常指相位，但仅在有些应用中才对应于化学意义上的相位），因此可在各个尖峰间设置阈值。如下所示，通过使用直方图或图像本身作为指导，也可自动调整阈值（Prewitt & Mendelsohn，1966；Weszka，1978；Otsu，1979；Kittler et al.，1985；Russ & Russ，1988a，2007；Rigaut，1988；Lee et al.，1990；Sahoo et al.，1988；Russ，1995c；Melgani，2006；Luessi et al.，2009）。在许多阈值级别，将先验知识与由图像中的特征得到的测量参数相比较的方法（Wolf，1991），更为专业。

在某些设置中，直方图中的尖峰对应于图像中的结构这一假设可能是对的，但对于许多真实图像却并非如此，如图 7.3 所示。在更为受控的光照条件下（如在天文学、遥感、平板扫描仪或实验室中所遇到的情形），更容易满足这一假设。这种假设特别适用于显微图像（包括有选择性染色或其他处理的

样本制备）。在某些复杂情形下，可直接对图像进行阈值处理，但通过检查直方图的各个尖峰来进行阈值设置是否可行，则取决于各个尖峰是否明显。在很多情形下，各个像素的亮度级并不唯一地与某个结构相关联。在有些这样的例子中，第 5 章中介绍的图像预处理运算，可变换图像中的原始亮度值来形成一幅新的图像，在这幅新图像中，像素亮度表示某些导出的参数，如局部纹理、亮度梯度或方向。

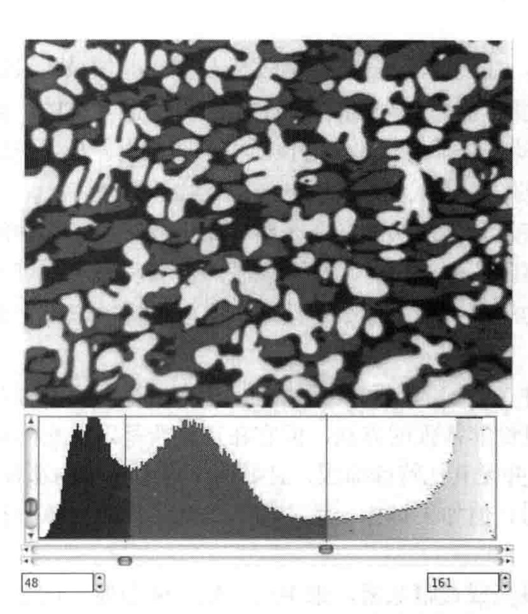


图 7.1　在直方图上设置亮度界限来对图像进行阈值处理的例子。鼠标用于拖动滑块，所选像素着色为彩色，并且显示了界限的数值

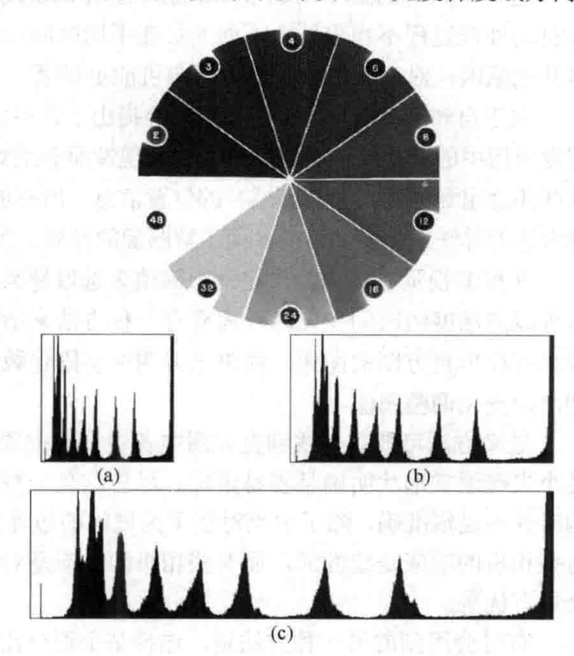


图 7.2　密度校准标准的 16 位图像及其亮度值的直方图：(a)传统的 256 通道直方图；(b)512 个通道；(c)1024 个通道

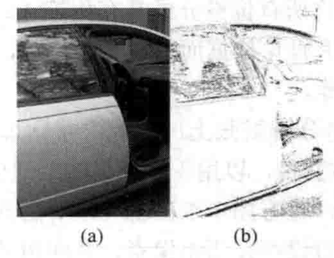
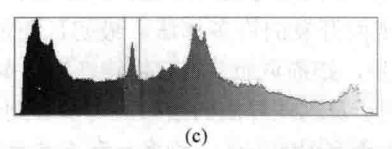



图 7.3　基于直方图中的尖峰对一幅真实图像进行阈值处理：(a)原始图像；(b)通过设置级别选取的像素；(c)显示了一个尖峰及阈值界限的直方图。具有所选亮度值的像素的位置并不对应于原始场景中的任何单个物体或结果单元

为进行测量，如求每个结构的面积比例，并非总是需要对图像进行阈值处理。直方图分析就可以完成这一任务，方法是使用高斯（或其他形状）函数来拟合直方图。给出可最佳拟合的高斯尖峰的数量，就可通过多重回归的方法求出位置、宽度和高度。这可以估算每相的面积，但在各个尖峰互相重叠时，则无法确定像素的空间位置。结果通常会很差，因为真实的成像情形很少产生理想的高斯尖峰（或任何其他一致的形状）。核磁共振图像是一个例外，在核磁共振中，这种方法一直用于生成如下图像：图像中重叠部分的像素并不转换成黑色或白色，而是根据两个重叠高斯尖峰对该亮度值的相对贡献转换成灰色（Frank et al.，1995）。当然，这不会产生常规意义上分割图像，但它可以产生描述重叠结构（如脑部扫描中的黑色和白色部分）的可视图像。

7.2 自动设定阈值

通常，人们通过手动调整阈值来产生看似正确的结果，但多数情形下应避免这样做。除了耗时且与自动处理过程不相容外，不同人员在不同时间可能会得到不同的结果。手动阈值处理误差相比于任何其他原因，对后续图像分析的影响可能更显著。

关于自动阈值处理过程，人们已经提出了许多算法，如下所示。主要用在工业质量控制领域的机器视觉应用中的有些算法，更注重的是重现性而非绝对精度。多数自动方法在其计算中会用到直方图，但其些还会用到像素（及其邻域）的位置信息。所有的自动方法都会对图像的性质做某些假设，且为了选择合适的算法，需要尽可能多地了解图像的性质、图像的获取方式以及面临的场景或物体类型。

在机器视觉系统中，设定常数阈值来选取感兴趣特征是不现实的。在光照、相机或待观察部件的位置或清洁度随时间变化时，需要有一种方法来适应实际图像。有时，控制光源的位置和颜色，可生成具有简单直方图的图像。如果直方图中有固定数量的清晰尖峰，那么自动跟踪尖峰位置可以为上面列出的变化调整阈值。

最简方法可能是，找到直方图中各尖峰的位置，并在它们的中间设置阈值。该方法很强大，因为各个尖峰通常形状明确且容易定位。尽管这是一种重复性非常强的方法，但它在定义所呈现的实际结构时并不是很准确。除了中点对应于区域间的边界外，并无其他特殊原因，且如第 1 章中指出的那样，有些相机的响应是线性的，而有些相机的响应是对数的。但如果目标一致，那么与此类似的简单方法会更有优势。

有时会用到的另一种方法是，选择某个百分比的最亮或最暗像素，来生成一幅二值图像。该过程是从直方图的一端开始，统计直方图各分段的和，直到到达期望的图像区域比例。然后把阈值设定为计数停止时最后一个分段的亮度值。在背光拍摄的工件中定位小孔或瓶上的标签等应用中，这种方法效果很好，因为小孔或标签在图像中的位置可能变化，但总面积保持不变。

基于直方图内容来计算阈值的一种常用方法是，将所有像素分成具有相等（或近似相等，具体取决于直方图中的有限计数）方差的两组。这种方法没有理论依据而只是一种经验，且要求一种选取了初值的迭代算法，来计算两组像素的方差、调整电平等。

为阈值自动设定而开发的许多算法，最初目的是分辨纸张上的印刷文字，即作为光学字符识别（OCR）程序的第一步，扫描页面并将它们转换为文本文件，以用于编辑目的或通信目的。图 7.4 显示了一页扫描文本的例子及这些算法的结果。可以看到，直方图中未出现“双峰之间为谷”的情形，因此不能使用查找双态直方图的方法。许多自动方法会假设存在两组像素，并应用统计检验来选取一个可将它们最好区分的阈值。

每种方法都会对直方图的性质做不同的假设，且可使用合适的统计或其他检验将直方图分成两部分，每部分表示出现的一种结构（纸和油墨）。在 Parker（1997）、Yager（1979）、Otsu（1979）和 Trussell（1979）中可找到一些更为广泛使用的方法，Sezgin & Sankur（2004）中详细介绍了所有可用的技术。下面仅介绍一些有代表性的方法。

Trussell 算法（Trussell，1979）可能是使用最广泛的自动方法，因为它通常可产生非常好的结果[见图 7.4(d)]且易于实现。它可找到一种阈值，以学生 t 统计的最大值生成两组（较亮像素组和较暗像素组），学生 t 统计由两组像素的均值（μ）、标准差（σ）和每组中的像素数量（n）计算得到，如式（7.1）所示。这是测量两组像素的不同概率的标准统计检验，因此找到使得 t 值最大的阈值设定，就可区分两组像素。但它隐含了一个假设，即像素组可由均值和标准差正确地表示，换言之，即假定它们是正态分布的，而且很多直方图中并没有高斯尖峰。

$$t = \frac{|\mu_F - \mu_B|}{\sqrt{\frac{\sigma_F^2}{n_F} - \frac{\sigma_B^2}{n_B}}} \tag{7.1}$$

可产生较好结果的另一种非参数（即不对直方图形状做出假设）方法，则使用两组像素（阈值之上和阈值之下）的熵，如图 7.4(e)所示。计算相关熵有几种不同的模型，因此产生的结果也会稍有不同（Kapur et al.，1985；Abutaleb，1989；Brink & Pendcock，1996）。式（7.2）给出了最简形式的计算，它通过选择一个阈值 t，使得前景和背景的总熵 H 最大。p_i 值的计算是实际直方图计数 $h(i)$除以该直方图部分中的像素总数。

$$H = H_B + H_F, \quad H_B = -\sum_{i=0}^{t} p_i \log(p_i), \ H_F = -\sum_{i=t+1}^{255} p_i \log(p_i), \quad p_i = \frac{h(i)}{\sum_k h(k)} \tag{7.2}$$

Paker 总结的许多算法对这幅图像产生了类似的结果，但有些结果中的字符未完全分开，有些结果中的字符却断开了。手写文字情形时，识别更为困难，因为线条的粗细变化很大（Solihin et al.，1996）。没有哪种方法能对所有类型的印刷品、纸张和图像采集设置有效。而且即使存在这样的方法，与仅包含两类结构（或更多结构）的各种图像相比，出现的问题会更专业。有些此类方法可推广到多组图像的情形。例如，学生 t 检验方法应用到两组以上像素时，就变成了方差分析法（ANOVA）。但通常而言，有必要具备所出现像素组数的先验知识。

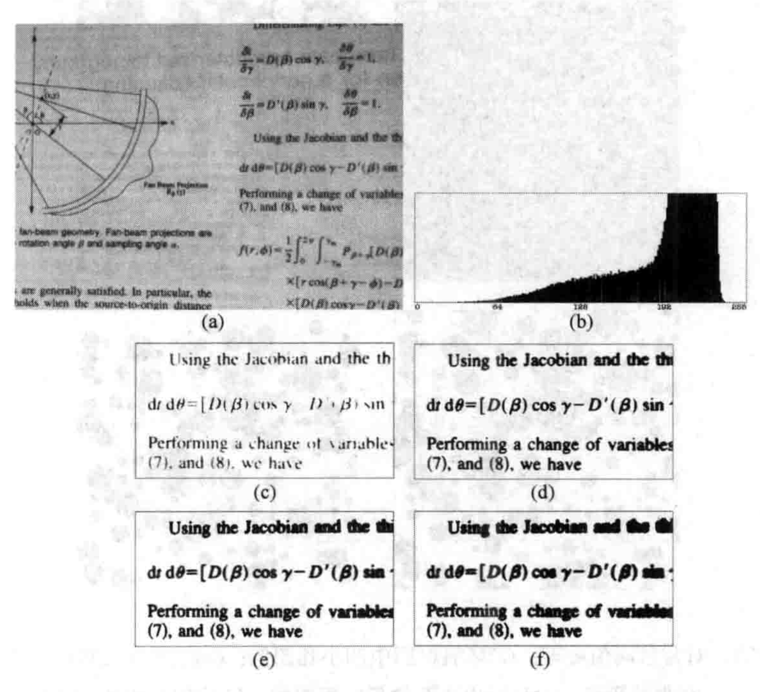


图 7.4　使用 Parker（1997）中的算法对纸张上的印刷文字进行自动阈值处理：(a)原始灰度扫描图像；(b)直方图，其中显示了下面几个例子中所用的设置；(c)Yager 算法，阈值为 134；(d)Trussell 算法，阈值为 172；(e)香农熵算法，阈值为 184；(f)Kittler 算法，阈值为 196

自动阈值处理的准确度和精度，取决于是否选择了合适的方法，这意味着要使用关于标本或场景和成像硬件的独立知识。Von Bradke（2004）声称，手动与自动将阈值设定到直方图中的最低点相比，结果的一致性更好（指可重复性而非准确度）。当然，选择最低点本身就是一种存在缺陷的方法，因为最低点易受噪声影响而且不稳定，当面积比例变化（完全不应发生）时会相对于峰移动。另外，改变

整个图像亮度（对比度、增益及伽马值）也会改变最低点。

Von Bradke 结果中最值得注意的问题是，所考虑的典型图像的直方图，要有一个大且明亮的尖峰（对应于抛光表面），且在孔隙对应的暗值端要有近乎平坦的拖尾。很多情况下直方图中并不存在清晰且明显的最小值，因此该技术基于计数中的统计波动来选择一个值，可以料想到结果中该阈值会剧烈地上下跳动。在观看图像的同时手动设定这些值时，操作者所用的算法是不同的（可能是基于细孔边缘平滑度的算法）。注意，此时能达到的可重复性并不意味着精度，即低图像放大率（许多细孔只有一个或几个像素大小），样品制备中边缘的平滑问题、穿透样品的电子束也可能在结果中产生偏差。自动方法的其他临界测试（如 Ozkaya et al.，2005；Francus & Pinard，2004；Jalba et al.，2004）也检查了特定算法对真实图像而非人造测试图像的性能。

本章后面将讨论一些不只考虑直方图形状的方法，如每个像素与相邻像素间的亮度差异，沿前景与背景区域边界的像素亮度值，或这些边界的形状。这些方法可为阈值处理选取更精确的阈值，但为选择合适的算法，需要关于图像性质的更多知识。

在许多情形下，如图 7.5 所示，感兴趣的物体既不具有单一的亮度值或颜色，也不在那些值的范围内。但是，如果物体周围的背景或物体所在平面的颜色或亮度非常均匀，那么通过阈值处理来进行选择是可行的。然后，将所得二值图像的对比度反转，就可得到物体的表示。在所示例子中，环氧基树脂中的细孔接下来必须使用第 8 章中介绍的分水岭技术来分开。

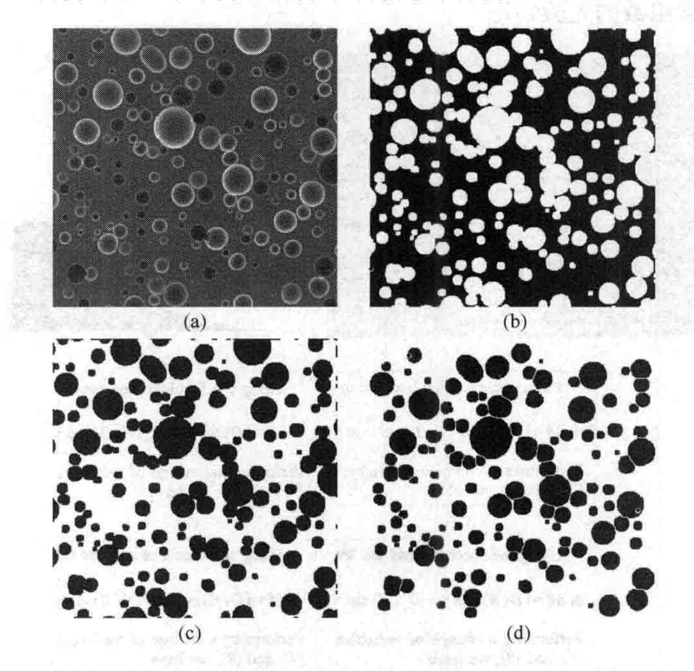


图 7.5　对背景阈值处理：(a)环氧树脂中的小孔图像；(b)经阈值处理后的中间灰度背景像素；(c)显示出小孔的反转后图像；(d)应用分水岭方法分开小孔

7.3　多波段图像

在有些情形下，可以使用同一场景的多幅图像来进行分割。最熟悉的例子是彩色成像，它使用不同波长的光。尤其对于卫星成像，它可包含几个红外波段，这些红外波段中具有根据植被、岩性等来选择区域的重要信息（Haralick & Dinstein，1975）。图 7.6 中给出了一个例子，其中每个波段的测量值

会因地形的不同而相互重叠，但组合两个通道后，它们完全分开了。

有些情形下，也可使用这种方法来对同一幅原始图像执行不同的处理运算，得到一系列图像。例如，将一幅包含亮度数据的图像与一幅包含局部纹理信息的图像合并，如下所示。

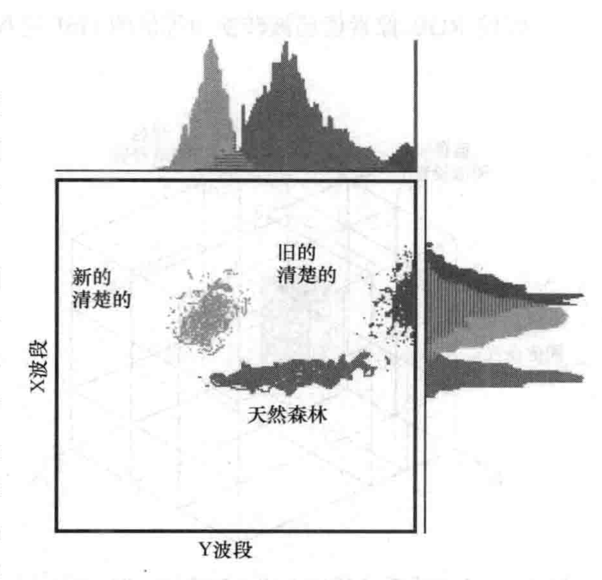


图 7.6　多波段卫星图像的地形分类示例。每个波段中各类间的重叠，要求相互重叠的两者均用于区分地形的类型

通常，独立的彩色波段或其他图像越多，分割效果就越好。在一幅图像中无法分辨的点，在另一个幅图像中或许可以完全分开。但对于多波段或多层图像，指定选择标准很困难。阈值处理的逻辑扩展是，对每幅图像设置亮度阈值，如指定红色、绿色、蓝色亮度的范围。然后，使用“与”运算合并这些准则（例如，如果某个像素的三个 RGB 分量都在所选范围内，就将它定义为前景的一部分）。这逻辑上等同于分别对每个图像通道进行分割，生成单独的二值图像，然后用布尔“与”运算将它们组合在一起。合并多幅二值图像的此类运算，将在第 8 章中介绍。

在单个处理中合并各个选择标准的原因是，帮助用户定义每个标准的范围。当各个颜色波段或其他多个图像亮度值单独设置时，最优设置及其相互作用并不明显。对一幅典型的可见光图像（多数情形下由 RGB 分量指定）进行阈值处理时，设计用户界面来选取某个特定的颜色范围，并不容易。为彩色图像设置阈值标准的一些方案正在使用之中。

这个问题包括几个方面。首先，尽管红色、绿色和蓝色灰度表示了检测器的工作方式、数据内部存储的方式，但它们并不对应于人们识别颜色或对颜色反应的方式。如第 1 章所述，人们更熟悉基于色调、饱和度和亮度或明度（HSI）的系统。如图 7.7 所示，有时仅用色调、饱和度或亮度通道之一也可得到满意的阈值处理效果，或如第 5 章所述，仅使用主分量通道也可获得较好的效果，但通常需要使用所有这些信息。每个 RGB 颜色通道的一系列直方图可能显示有尖峰，但用户通常并不能够判断哪些尖峰对应于感兴趣的各个特征。

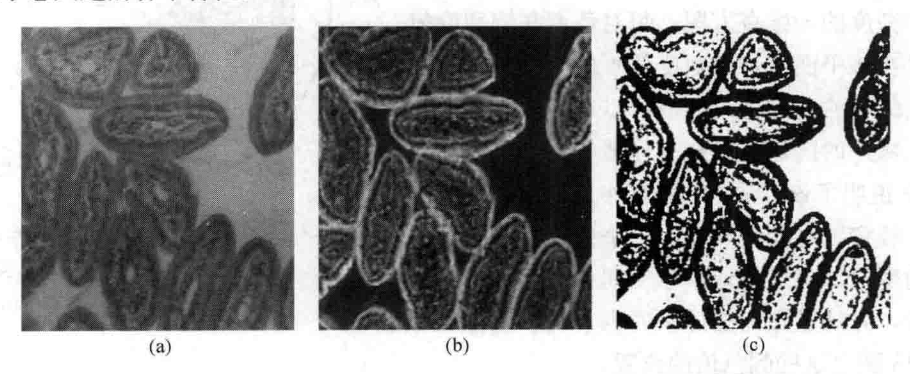


图 7.7　使用单个通道对彩色图像进行阈值处理：(a)原始染色生物薄切片；(b)由存储的红色、绿色和蓝色值计算的灰度值；(c)色调图像经阈值处理后显示了着色结构

即使 RGB 像素值已被转换为等价的 HSI 值且在该空间中构建了直方图，使用三个单独的直方图和三组阈值电平仍然无助于用户看到出现了值的哪些组合，或者哪些像素有那些组合。

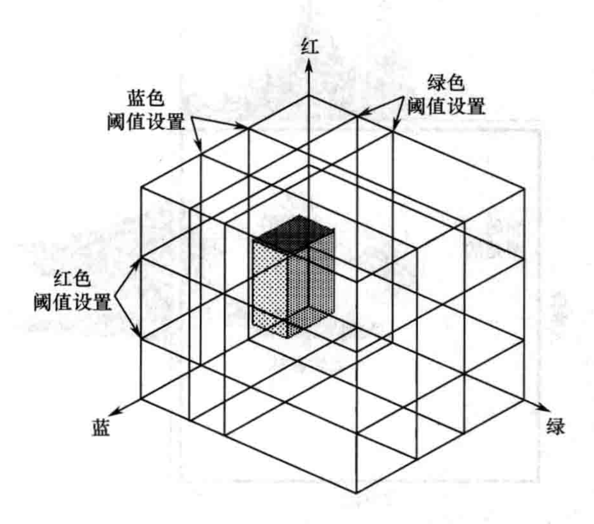


图 7.8 组合各个颜色通道上的单独阈值示例。阴影区域是红、绿和蓝的三个阈值设置的“与”运算结果。在三维空间中可以形成的唯一形状是长方体

对于三维彩色空间（RGB 或 HSI），交互式阈值处理比图 7.6 中所示的二维情况更为困难。对于显示或控制目的，不存在容易或明显的方法来交互式地圈出三维空间中的任意一个区域，并观察选中了哪些像素，或调整该区域来观察对图像的影响。在图像中标记一个像素或区域，然后观察在色彩空间中直接标记的颜色值（RGB 或 HSI），是有帮助的。对于三种以上颜色的情况（如卫星多波段图像），情况甚至更为复杂。

使用三个一维直方图和三组阈值电平（如 RGB 情形），并用逻辑“与”运算组合三个标准，可选择简单棱柱状彩色空间中某部分内的像素，如图 7.8 所示。如果色彩值的实际分布在彩色空间中具有某种其他形状，例如在不平行于轴的方向拉长，那么这个简单的长方体将不足以描述想要的颜色范围。

7.4 二维阈值

使用二维标准可设置更为灵活的阈值，方法是在任何彩色空间（RGB、HSI 等）进行，但用 RGB 坐标很难解释这种设置的意义。这也是反对为彩色图像使用 RGB 空间的许多争论之一。但这种方法非常适合于由色调和饱和度编码的彩色图像。HS 平面可表示为一个圆，其中方向（角度）与色调成正比，半径与饱和度成正比（见图 7.9）。图像的灰度或明度垂直于这个平面，并需要另一维来显示或控制。

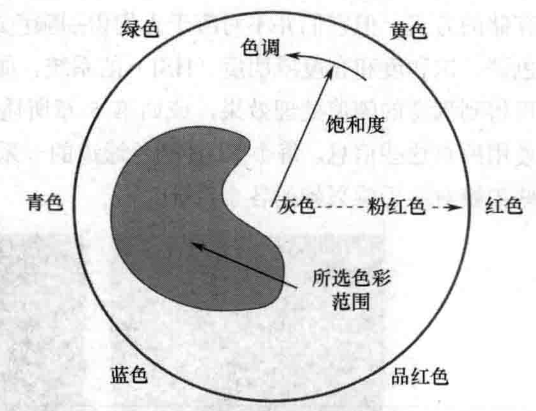


图 7.9 在二维参数空间（此处是色调/饱和度圆）中任意选择一个区域来定义将被选择用来进行阈值处理的颜色的组合原理

该图形表示了 HS 平面上的一个二维显示，而不是黑白图像中亮度的一维直方图。每对色调和饱和度值的像素数量在该平面上可以灰度值画出，表示一个具有多个暗色尖峰的直方图。阈值可以选为除简单、凸状、连通区域外的区域，这样更能适应实际数据的分布。图 7.10 说明了该方法，注意其中选择了较小范围的色调值、较宽范围的饱和度值，来处理由每颗圆形糖果导致的变化。也可在该彩色空间中找到一些位置来作为参考，方法是指向图像中的像素，以便程序可突出显示 HS 圆上这些颜色值的位置。

如图 7.5 中的例子所示，与糖果相比，背景在颜色和亮度上更均匀。这些像素位于 HS 圆的中心附件，因为它们的饱和度很低。如图 7.11 所示，对它们进行阈值处理产生了所有糖果的一幅二值图像。

利用第 8 章解释的分水岭分割算法分开接触的物体，并将这幅图像与原始图像组合（如保留更亮的像素）后，恢复了这些糖果的颜色。

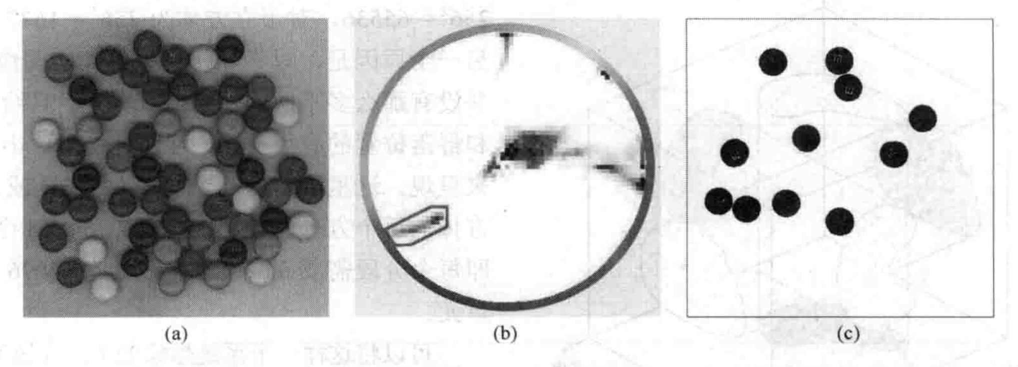


图 7.10　色调－饱和度选择示例：(a)原始图像；(b)从直方图中为单颗糖果选择色调－饱和度值（它作为灰度值显示在色调－饱和度圆上，表示像素数量；注意每种颜色群包含有靠近圆心处的低饱和度灰色背景）；(c)所得二值图像

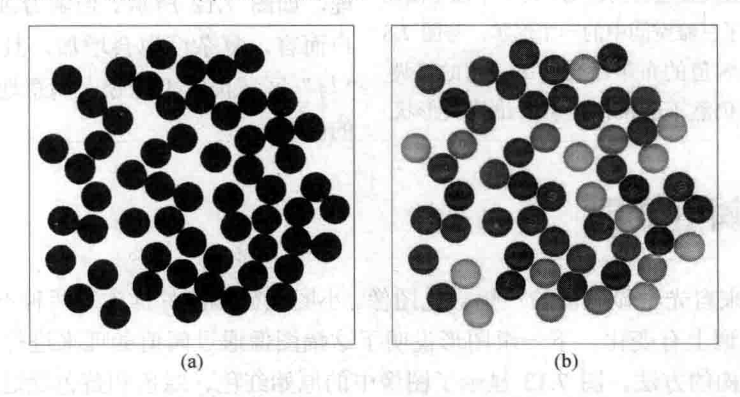


图 7.11　对一幅彩色图像中的背景进行阈值处理［始于图 7.10(a)］：(a)阈值处理并反转对比度后的图像中勾画出了糖果；(b)经分水岭分割并将二值图像与原始图像组合后，使得像素值更亮，分开的糖果可用于测量或计数

使用其他通道和坐标，可实现类似的直方图显示和阈值设置。对于彩色图像而言，HS 平面有时会显示为六边形（角点分别是红色、黄色、绿色、青色、蓝色和品红色），而不是圆形。第 1 章中的 CIE 颜色图也可用于此目的。对于有些卫星图像，近红外和远红外线亮度形成一个平面，在这个平面中可显示和选择热红外线（IR）和反射的红外线。

HS 平面可以画为表示 HSI 空间的立方体的一个方形面。对于计算机图形学显示，这更简单，并用到了接下来的几个例子中。

但是，具有方形面的 HSI 立方体拓扑上不同于第 1 章中用于表示 HSI 空间的圆锥体或双锥体，且方形 HS 平面拓扑上不同于图 7.10 中的圆。在这个方形中，最小色度值和最大色调值（橘红色的和品红色）隔得很远，而在圆中，色调是一个环绕的连续函数。这就使得使用方形来进行阈值处理有点不直观，但在多数情形下它仍然优于使用 RGB 彩色空间。

对于二维方形图，坐标轴可能有我们不熟悉的含义，但是基于值的组合来显示点的直方图的能力，以及基于直方图来选择阈值边界的能力，与多个一维直方图和多个阈值相比，优势明显，即使它并不容易推广到 n 维情形。

通常，直方图数组的大小一定程度上可因存储图像的各种 RGB 或 HSI 值的实际分辨率而减小。这不仅是因为数组的尺寸会非常大（对于方形为 $256^2 = 65536$，对于立方体为 $256^3 = 16777216$）。另一个原因是，对于一幅典型的真实图像来说，并没有那么多不同的值对或三元组，且给出尖峰和群落位置的有用显示，可通过使用更小的分段来呈现。这里给出的几个例子对 RGB 或 HSI 立方体的每个方形面使用 32×32 或 64×64 个分段，即每个分段需要 $32^2 = 1024$ 或 $64^2 = 4096$ 个存储单元。

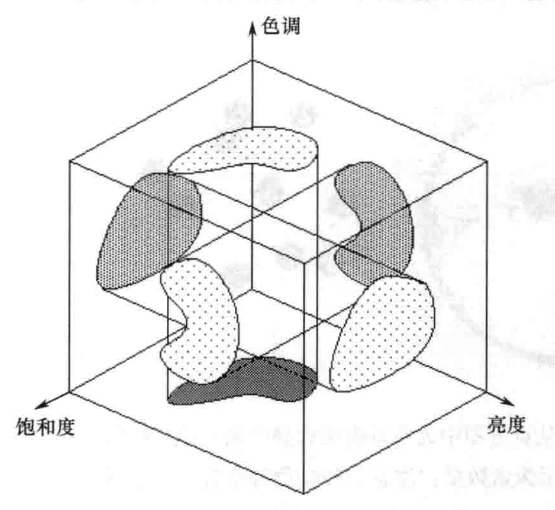


图 7.12　两个参数阈值设置组合的说明。每个平面中圈出的区域定义了三维空间中的一个形状，与图 7.8 中简单一维阈值的布尔组合相比，它的可调性更强，但仍然不能模拟任意三维聚类形状

可以将这样一个系统想象如下：在该系统中，由每对信号定义的每个二维平面，用于绘制轮廓阈值，然后将这些轮廓投影回多维空间而定义阈值处理，如图 7.12 所示。但随着维数的增加，对于用户而言，复杂度也会增加，且由多个投影定义的“与”区域仍然不能令人满意地拟合不规则或倾斜的区域。

7.5　多频段阈值处理

图 7.7 显示了来自光学显微镜的一幅彩色图像。小肠显微镜薄片标本用两种不同的颜色着色，因此阴影、色彩和色调上有变化。下一组图形说明了这幅图像通过阈值处理来进行分割，进而使用该信息来隔离特殊结构的方法。图 7.13 显示了图像中的原始红色、绿色和蓝色通道，以及它们各自的亮度直方图。图 7.14 显示了投影到 RGB 彩色立方体的红/绿、绿/蓝和蓝/红面上的像素值直方图。注意，图像中的多数像素在所有表面上都有沿着立方体对角轴聚集的趋势。换言之，对于多数像素来说，这种向某种颜色靠近的趋势，通常会因增大所有颜色值而增大亮度。这就意味着 RGB 空间很难使得不同彩色值分散开，并且不能方便地设置阈值来区分所呈现的不同区域。

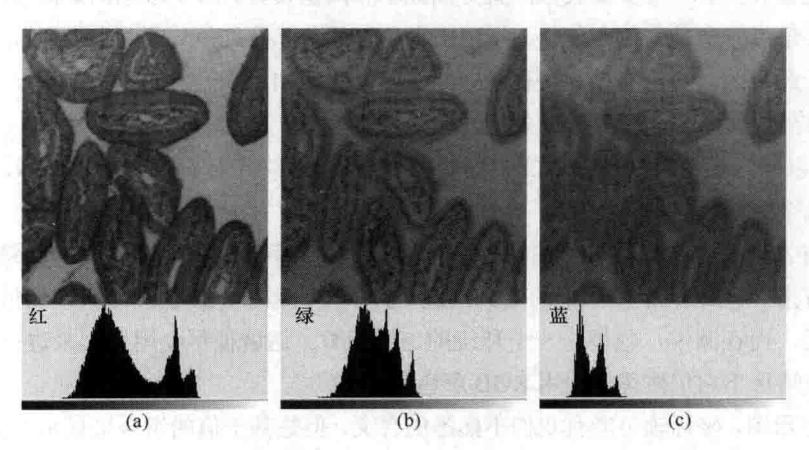


图 7.13　图 7.7 中图像的红色、绿色和蓝色通道，以及它们的亮度直方图

图 7.15 显示了图 7.13 中颜色信息至色相、饱和度和亮度图像的转换，以及这些通道各自的亮度直方

图。图 7.16 显示了投影到各自的二维色相/饱和度、饱和度/亮度和亮度/色相方形面上的值。注意各直方图中分散得很开的尖峰使用更多色彩空间并分隔几个不同群落值的方式。一般来说，对于生物样本上的染色剂而言，色相图像识别特定染色的位置，而饱和度图像对应于染色量，亮度图像对应于染色标本的总体密度。

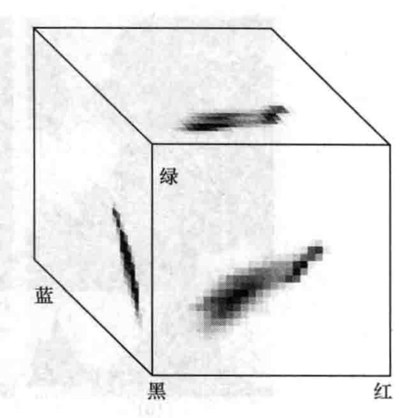


图 7.14 图 7.12 中图像的像素值对，在 RG、BG 和 RB 平面上画出，并投影到立方体的表面上

第 5 章介绍的图像主成分分析（PCA），也会生成这样一些图像，在这些图像中，各个派生通道有更分散的值，这些值通常比原始的红色、绿色和蓝色通道更适合于阈值处理。主成分分析的一个优点是，轴是不相关的，而且沿每个平面中的每个轴选取一个范围，确实可以高维体积来正确地描述体积。图 7.17 显示了主成分通道及它们的直方图，图 7.18 显示了二维图。第 5 章中的图 5.57(b)以三个任意分配给红色、绿色和蓝色的主成分通道，显示了这幅同样的图像。

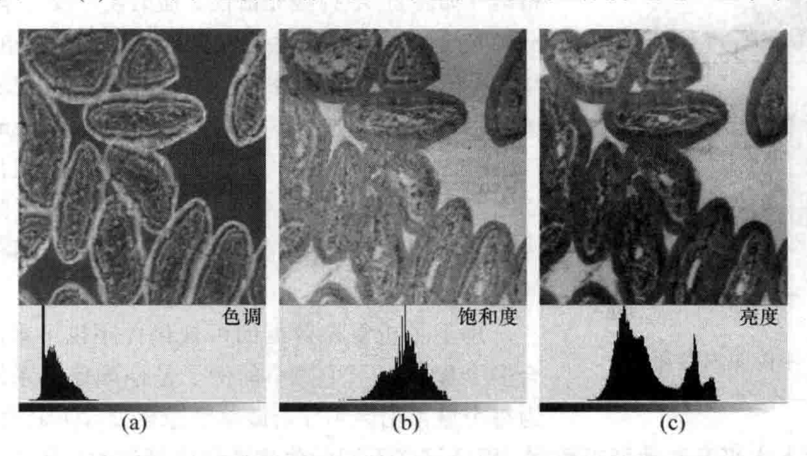


图 7.15 图 7.7 中图像的色调、饱和度和亮度通道，以及它们的亮度直方图

多频段图像并不总是具有不同的颜色。一个常见的例子是 SEM（扫描式电子显微镜）多元素 X 射线图，组合这些图像并基于成分，可选取感兴趣的相位。在许多情形下，可通过对每幅图像分别进行阈值处理，然后应用布尔逻辑来组合图像。当然，可能首先需要对噪声非常大的原始 X 射线图像进行处理（如使用中值滤波来消除噪声），以降低像素间的统计差异（见第 4 章），然后进行二值图像处理（见第 8 章）。

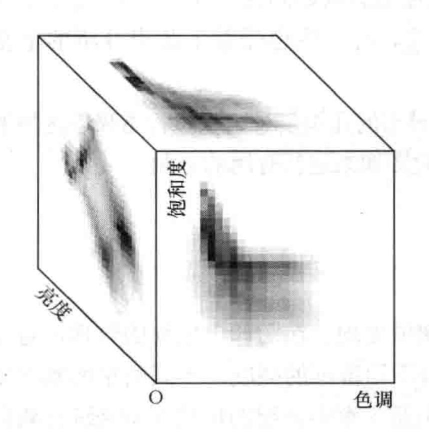


图 7.16 图 7.15 中图像的像素值对，在 HS、SI 和 HI 平面上画出，并投影到一个立方体的表面上

利用 X 射线或其他特定于元素的信号，如二次离子或俄歇电子，基本上可检测出整个元素周期表内的元素。因此，可指定非常复杂的元素（这些元素必须是存在的或缺失的）组合，或可将所需近似亮度级（因为亮度大体上与元素浓度成正比）用来指明感兴趣区域。对这些元素图像的组合进行阈值处理，就生成了有时称为化学图的结果。当然，几种元素可能会在（如金属、矿物质或生物组织）标本的同一区域内出现的这一事实，并不能直接表明它们是化学组合的。

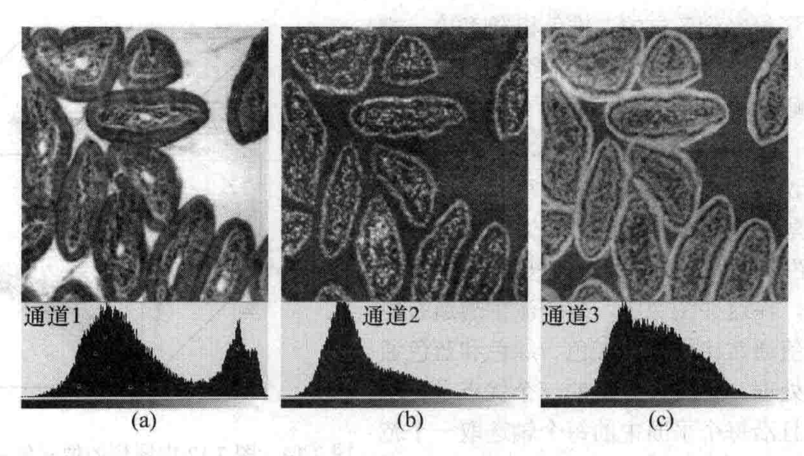


图 7.17　图 7.7 中图像的各个主成分通道及它们的亮度直方图

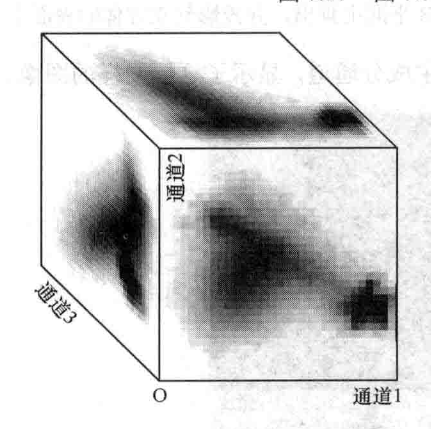


图 7.18　图 7.17 中图像的像素值对，投影到一个立方体的表面上

原理上讲，我们可为图像中的每个像素存储一个完整的分析谱（如源自 X 射线光谱仪、质谱仪，或光学可见光或红外线分光光度计），然后使用合适的计算导出每个点处的实际组成信息，或者运用主成分分析找出谱中最重要的数据，并最终将这些信息用于阈值处理来选择感兴趣的区域（Watanabe et al.，2009；Parish &Brewer，2010；Booth et al.，2010）。目前，这种方法由于需要大量的存储空间和冗长的计算而应用受限。随着更快、更大型的计算机和存储设备变得普遍，这类方法将会得到了更广泛的应用。

用于分析复杂数据的可视化程序也可采用布尔逻辑来组合多个参数。一个简单的例子是地理信息系统（GIS），其中为每个城市街区（可当做单个像素处理）记录多种数据，如人口密度、平均收入水平及其他调查数据。组合这些不同的值来选择区域进而试销商业产品是一种标准的商业手法。另一个例子是显示金属梁受载荷时计算出张量性能，就如同在一个计算机程序中建模一样。复杂动态系统（如逐步形成的暴雨）的超级计算机仿真，可产生也受益于此类分析的丰富数据。

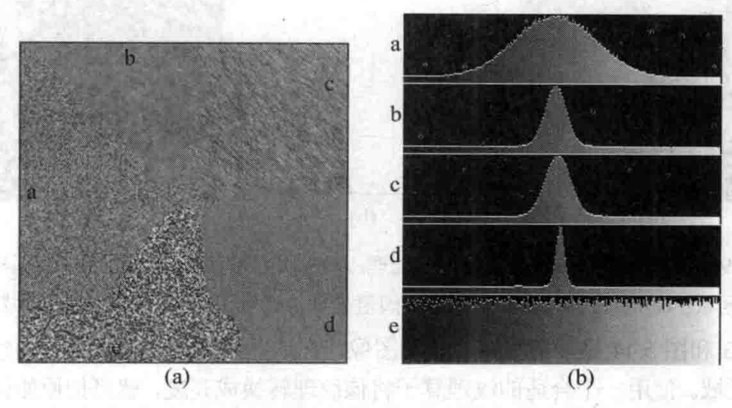
图像处理还有一些其他用途，它可通过由单幅原始灰度图像导出的几类信息，来执行感兴趣区域的选择性阈值处理。该处理可生成其他图像，这些图像可当做多频段图像来进行有用的分割。

7.6　来自纹理的阈值处理

仅使用原始亮度值，几乎不能对实际图像进行令人满意的阈值处理。在图像中出现的纹理信息是最强大的可用工具之一。我们可能会遇到几类不同的纹理，包括不同范围的亮度、不同的空间频率以及不同的方向（Haralick et al.，1973）。下面几张图片显示了使用第 5 章中介绍的那些变量来进行阈值处理的方法。

图 7.19 显示了一幅测试图像，其中包含了视觉上可由纹理区分的 5 个不规则区域。如亮度直方图所示，每个区域的平均亮度是一样的。区域(e)中包含的像素具有均匀的随机亮度值，亮度的值域为 0～255。从区域(a)到区域(d)具有高斯亮度变化，对区域(a)和区域(d)来说，像素位置也是随机指定的。对

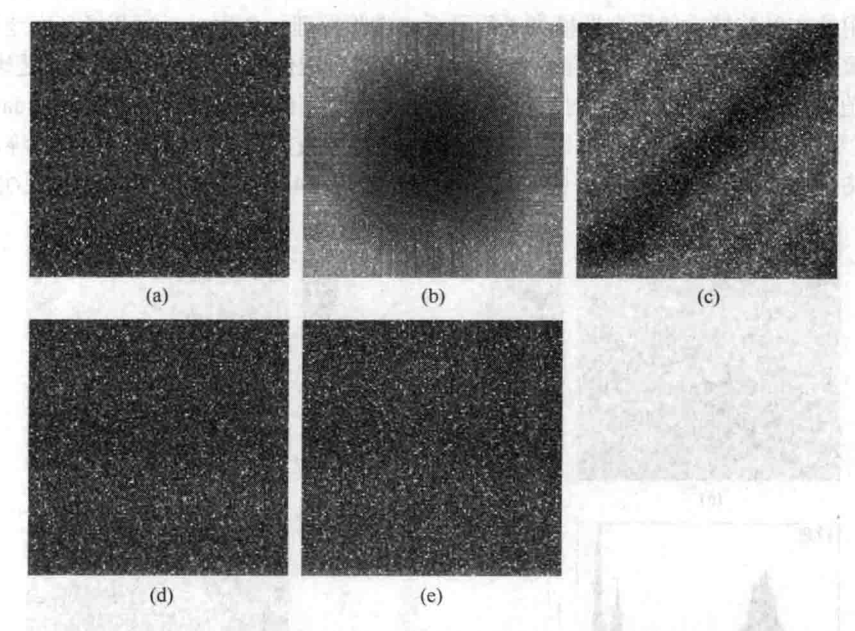
区域(b)而言，这些值已用高斯平滑进行了空间平均，因而减少了变化量。对于区域(c)，像素已在某个方向平均，以便创建一种方向性纹理。



(a) (b)

图 7.19　测试图像，其中包含了将由纹理差异分辨的 5 个不同区域。亮度直方图如图(b)所示，每个区域的平均亮度是相同的

针对纹理特征，一种常用（且有时有用）的方法是第 6 章中介绍的傅里叶变换。图 7.20 显示了图 7.19 中每种模式的功率谱。区域(b)中的平滑相当于一个低通滤波器，它削弱了高频部分。区域(c)中的方向性在频率变换图像中是可见的。对于其他区域，随机像素分配并不会在频率域中创建任何不同的模式。此时，它们不能用于选取不同的区域。


(a) (b) (c)

(d) (e)

图 7.20　图 7.19(a)中每个区域内图案的二维 FFT 功率谱。尽管可看到一些细微差别［如区域(b)中缺少了高频，而区域(c)中缺少了方向］，但它们不能生成令人满意的分割

对这幅图像，计算出移动半径为 4 个像素的邻域中的方差，作为图 7.21 中的空间域纹理敏感算子。第 5 章中描述的这一操作，会生成一幅导数图像，该图像中的每个区域会有唯一的灰度值，如亮度直方图中所示。由于 5 个尖峰是分开的，因此可直接进行阈值处理。该图中显示了已圈出所选区域（以不同的颜色表示）的一幅图像，它是通过对方差图像进行阈值处理后，叠加到原始图像上形成的。因为纹理的空间尺度为几个像素宽，因此区域的边界位置必然有几个像素的不确定。同理，视觉上很难估计精确的边界位置。

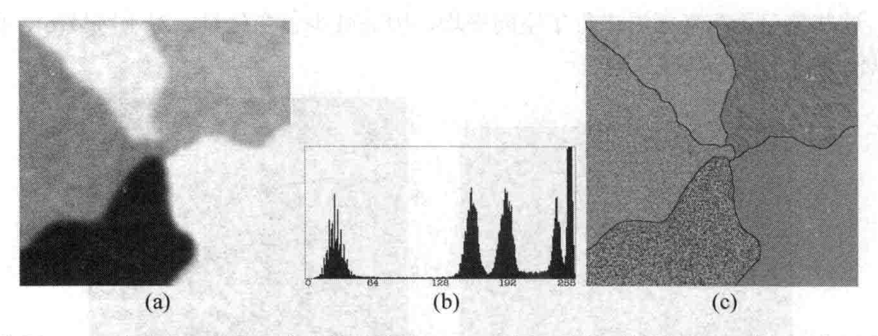


图 7.21　将图 7.19(a)中的纹理转换成亮度差异的处理：(a)应用方差算子的结果；(b)直方图表明有 5 个亮度不同的区域；(c)对图(b)中的直方图进行阈值处理，选取了每个区域，这些区域由不同颜色显示

第 5 章中图 5.43 和图 5.44 显示了很多典型的图像，在这些图像中，视觉上可由纹理的不同来区分没有明显亮度差异的区域。使用一个合适的纹理算子将该纹理转换成亮度，就可把原始图像转换为可进行阈值处理的图像。第 5 章中给出了用于此目的的几个纹理算子（最大－最小算子、分形维数算子）。

7.7　多个阈值处理标准

图 7.22 显示了一幅稍显复杂的测试图像，其中有些区域可由不同的空间纹理区分，有些区域可由不同的平均亮度区分。不存在可用于区分所有 4 个区域的单一参数。应用方差算子就可提取出的纹理，可产生一种有用的灰度差异。使用高斯模糊进行平滑来消除纹理，会生成一幅凸显亮度差异的图像。必须使用两幅图像来选择各个区域，方法是对每个区域单独进行阈值处理，然后使用布尔逻辑（见第 8 章）合并这两幅二值图像。另一种方法是，对彩色图像使用上面介绍的同类二维直方图（Panda & Rosenfeld, 1978）。图中显示了各幅图像的直方图和二维直方图。在每幅直方图中，均只有 3 个尖峰，原因是这些区域在亮度或方差上，都不是完全分开的。在二维直方图中，4 个区域都有各自的可见尖峰。

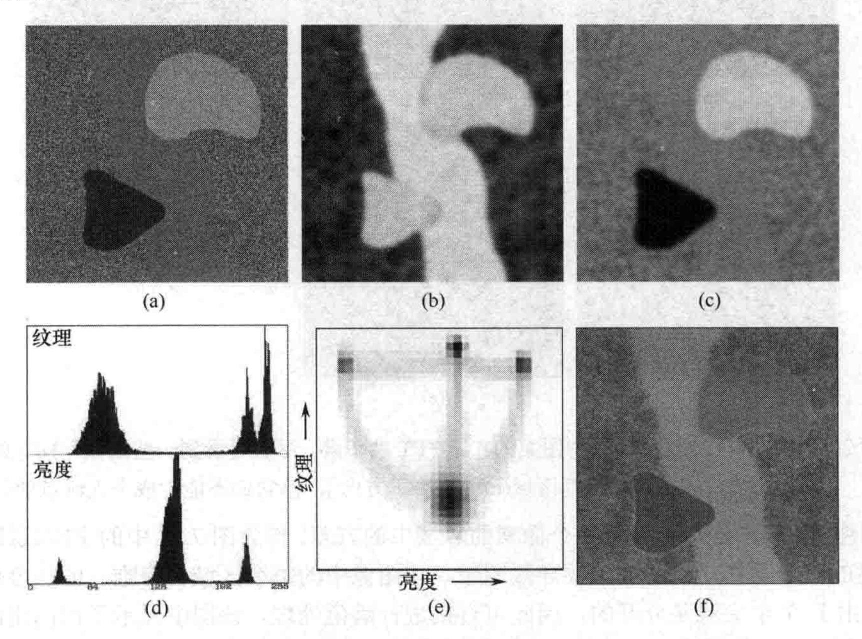


图 7.22　使用多个标准分割：(a)原始图像；(b)方差算子生成的纹理图像；(c)平滑生成的亮度图像；(d)图(b)和图(c)的直方图；(e)二维直方图；(f)正文中描述的彩色合成图像

这种情形类似于使用不同的颜色通道，如上面的图 7.6 所示。实际上，可成功分割图像的不同导数图像，有时会使用不同的颜色通道来显示。这纯粹是一种视觉效果，因为所表示的数据与颜色无关。但它确实利用了人类视觉能很好地分辨颜色这一事实，而且它还揭示了这个基于多个纹理和亮度标准的阈值处理的例子，与基于各个颜色通道的彩色图像阈值处理的例子之间的相似性。图 7.22(f)显示了一幅 Lab 彩色图像的信息，其中 L 通道中的信息是原始图像，a（红－绿）通道中的信息是平滑后的亮度值，b（黄－蓝）通道中的信息是来自方差算子的纹理信息。

图 7.23 给出了使用两个标准对食品冰晶图进行阈值处理的例子，其中一个标准就是纹理。分形纹理算子部分地勾画出了冰晶，就像色相通道的作用那样。分别进行阈值处理，然后使用布尔逻辑（见第 8 章）合并合并，就产生了一种有用的冰晶分割，如图中的轮廓所示。

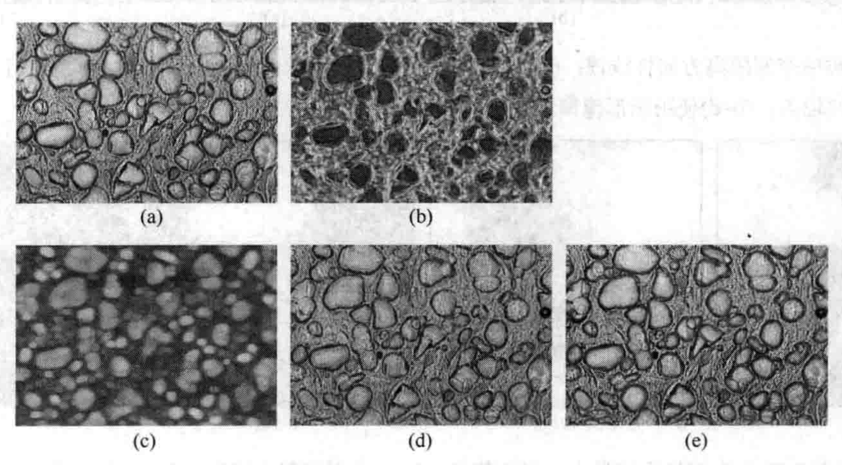


图 7.23　食品冰晶的分割：(a)原始图像；(b)对亮度通道应用分形纹理算子的结果；(c)色相通道；(d)在 Lab 图像中，L 通道中的原始亮度、a 通道中的纹理以及 b 通道中的色相形成的彩色合成图像；(e)经阈值处理和合并的最终结果在原始图像中显示为轮廓

7.8　纹理方向

图 7.24 显示了另一幅测试图像，其中包含具有相同均值亮度、亮度分布，以及局部变化但纹理方向不同的空间尺度区域。在二维频率变换中这很明显，如图 7.25(a)所示。变换中的三个辐条揭示出了三个范围的空间域方向。

先使用具有平滑边缘（见第 6 章）的一个扇形区域掩模来选取每个辐条，然后反变换图像生成图 7.25 所示的三幅空间域图像。原始图像中每个纹理方向都是独立的，在其他位置具有均匀的灰度背景。但由于这些纹理区域中的亮度值覆盖了周围的一些区域，因此不能通过对这些图像进行阈值处理来勾画出区域。

图 7.24　包含具有不同纹理方向但平均亮度、标准差和空间尺度相同的区域的一幅图像

对一个 5×5 像素的八边形邻域使用一个最大－最小算子，如图 7.26 所示，抑制了均匀的背景区域，同时突出各个纹理区域。对这些图像进行阈值处理，然后应用闭运算（见第 8 章）填充内部的间隙并平滑边界，就生成了每个区域的图像。图 7.26(d)显示了合成的结果。注意，对图像的边框处理效果并不好，原因在于傅里叶变换

是把边缘当做连续来处理的，详见第 6 章中的讨论。此外，各个区域的边界是不规则的，这一结果中只是近似地勾画出了边框。

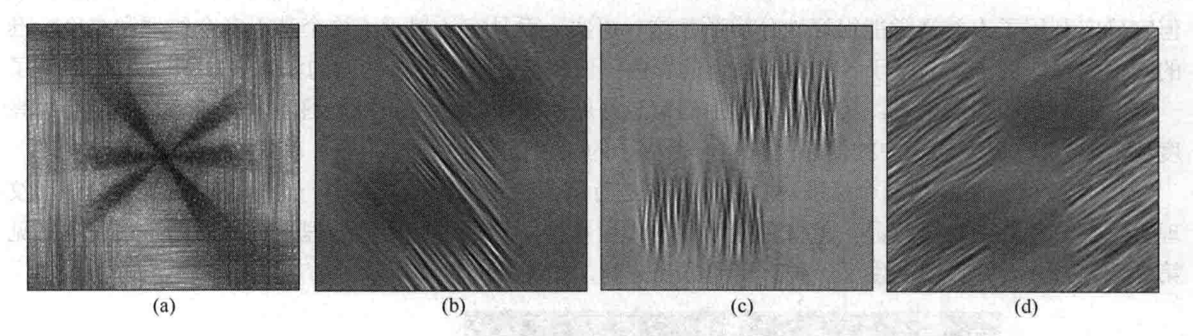


图 7.25 在频率空间隔离方向性纹理：(a)图 7.24 中图像的二维频率变换，其中显示了对应于每种纹理对齐的辐条；(b~d)使用扇形掩模选取每个方向并反变换的结果

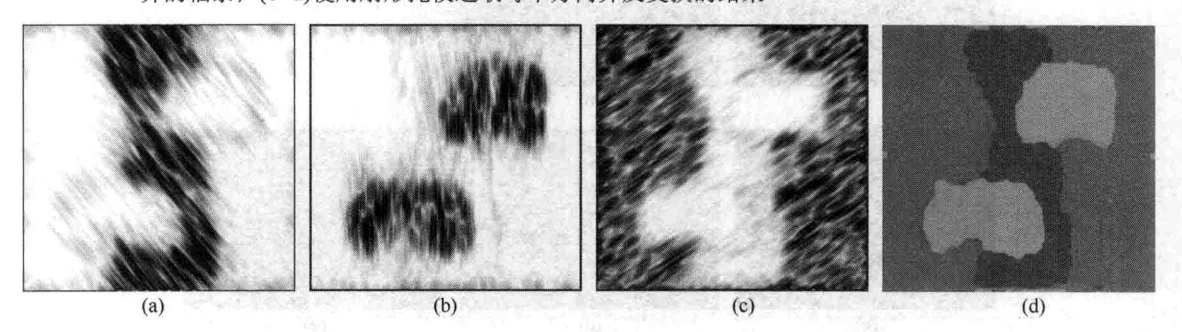


图 7.26 对图 7.25 所示图像应用最大－最小算子：(a~c)分别对图 7.25(b)、(c)和(d)进行处理后的结果；(d)对这些图像阈值处理后，所选区域的彩色组合

多数情形下，对纹理方向首选空间域处理方式。图 7.27(a)显示的是对一幅图像应用 Sobel 算子后的结果，如第 5 章中所述。使用 3×3 邻域算子得到了 x 和 y 方向的两个方向性一阶导数。接下来使用反正切函数来组合这两个一阶导数，得到一个角度，即最大亮度梯度方向。然后将所得角度标定到图像的亮度范围 0～255，使得每个亮度阶对应于约 1.4°。

图 7.27(b)所示的亮度直方图中有 6 个尖峰。它们以相隔 180°（128 个灰度值）的方式成对出现，因为在每个纹理区域中，最大梯度的方向可能是两个相反方向中的任一个。将整个灰度值的范围指定到 0°～180° 及 180°～360°，可将角度值减少到 3 个方向［见图 7.27(c)］。图 7.27(d)和图 7.27(e)显示了应用一个中值滤波器之后的结果。

对一幅由灰度表示方向的图像应用任何邻域处理运算（如平滑或中值滤波）时，需要有特殊规则来说明在 0 处值的模的变化。例如，包含像素值 15 和 251 的邻域的平均值是 5，而不是 133。达此目的的一种方法是，对每个邻域处理两次，一次是处理存储值，另一次是将原始像素值移位为“(值＋128) 模 255”，然后保存更小的那个结果。

图 7.27(f)显示了三次阈值处理运算来选取各个纹理方向的结果，每个结果都被指定到了一个颜色通道。沿边界出现的一些黑色像素未被这个自动算法选取，但它们是相对次要的，因此并不完美地定义出了原始图像中的区域边界。这个结果优于频率变换法，而且有更平滑的边界，与视觉判断结果的吻合度也更高。

图 7.28 说明了对一幅真实图像使用这一过程的方法。人字形布料有两个主要的方向区域，这两个区域在颜色和亮度上都不可区分。Sobel 方向算子指定了表示方向的灰度值，如第 5 章中的图 5.45 所述。对结果图像进行阈值处理，就勾画出了原始图像中的区域。

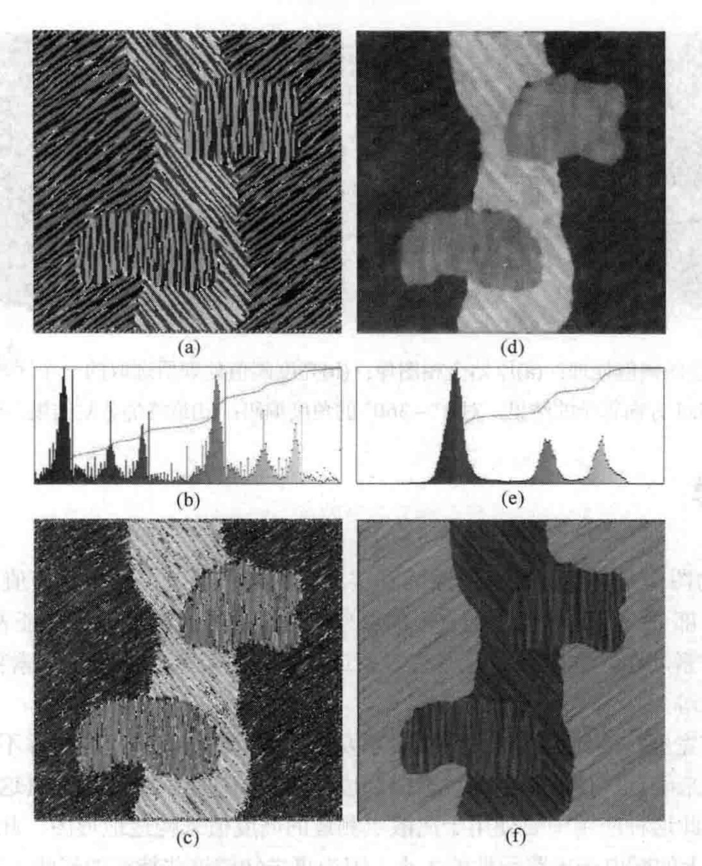


图 7.27 使用 Sobel 方向算子隔离方向性纹理：(a)将 0° ~ 360° 范围内的角度值指定为唯一的灰度值后，图 7.24(a)中图像的结果；(b)图(a)的直方图，它显示了 6 个尖峰；(c) 将 0° ~ 180° 和180° ~ 360° 范围内的角度值指定为相同灰度级值后的结果；(d)对图(c)应用中值滤波后的结果；(e)图(d)的直方图，它显示了 3 个尖峰；(f)基于图(d)，对每个区域进行阈值处理，然后对每个二值结果指定一个颜色通道，并将它们叠加到原始图像上后的结果

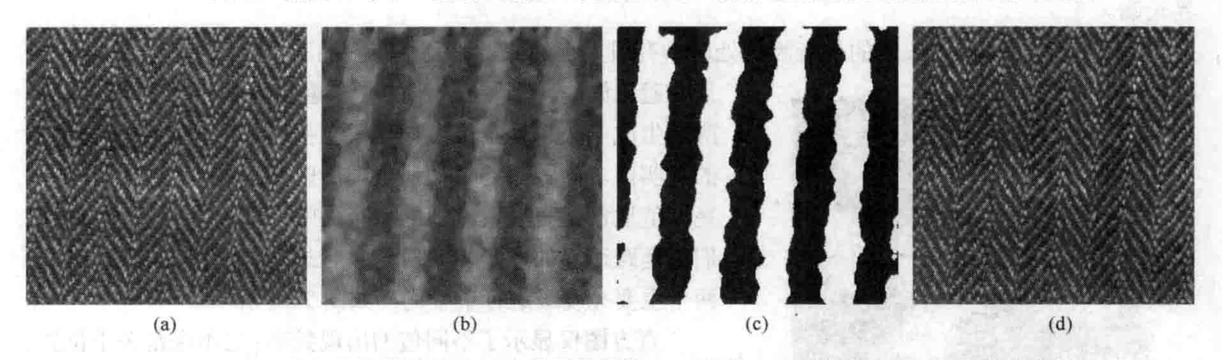


图 7.28 基于方向的阈值处理：(a)布料原始图像；(b)应用 Sobel 算子对 0° ~ 180° 的角度编码；(c)阈值处理后的结果；(d)将二值图像叠加到原始图像上，显示了所选的区域

图 7.29 显示了另一个合成物的例子，即带有薄片结构的金相样品。将其分割成各个区域需要几个步骤。亮度阈值处理能直接描述无纹理区域。运用 Sobel 方向算子能生成一幅可被阈值处理来勾画出其他两个区域的图像。结果是一幅完整的样品图，如图所示。

涉及亮度（通常在不同的颜色通道内）和纹理的这些合成方法，尤其适合于识别卫星图像中的农作物。不同种植模式会产生不同的纹理，使用可见光和红外通道组合这些纹理，可产生非常具体的结果。

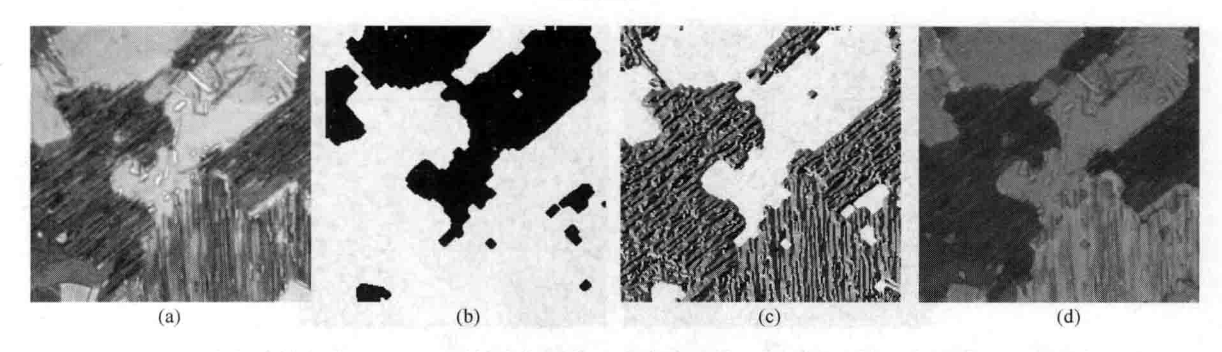


图 7.29 多个标准的阈值处理：(a)原始金相图像；(b)亮度阈值处理所选取的一个区域；(c)对剩余区域应用 Sobel 方向算子的结果，对 0°～360° 的角度编码；(d)最终的分割结果，所有区域均已着色

7.9 区域边界

由于基于像素的图像是对正被表示的连续真实场景的最佳逼近，且由于阈值处理会将每个像素分类为前景或背景的一部分，因此只能达到有限的精度。基于边界线的另一种特征表示方法可能更精确。相较于比较粗糙的像素间距，它们可以是具有多条边和多个角点（定义为子像素精度的 x, y 坐标）的多边形，或样条曲线等。

由于线本身没有宽度，因此这样的边界线表示要优于精确测量。但确定线并不容易。各个点的位置可通过在像素间内插来确定，使用数学函数来拟合边界两侧的像素，或许可改进这一结果。此类方法常用于地理应用中，因此这种应用中要使用于离散点测量的高度值来构造地形图。此类方法也用于计量应用中，比如测量晶圆上的微电子电路元件的大小，因为事先知道这些特征的形状（通常由直线组成），因此这种测量是可能的。这类专门应用超出了本章中介绍的典型图像处理运算范围。

采用内插技术来平滑跨越像素边界线的一种方法，已被第 10 章中的超分辨率周长测量程序使用，目的是进行特征测量。这种方法使用邻域处理（第 5 章提及的高斯－拉普拉斯算子或 LoG 算子）来拟合通过每个像素的一条自适应边界线，进而提升精度和分数像素准确度。下面会进一步讨论内插的边界轮廓。

基于产生图像像素表示的传统阈值处理，有限大小的像素只允许有限的准确度。但由于需要在结果中避免偏差，对同一场景的许多重复图像上执行相同运算，应会生成一个平均结果，该结果会接近尺寸或其他特征测量的真实值。问题始终是如何处理沿区域边界的像素，它们的亮度值通常在亮度直方图的各个尖峰之间。多数情形下，它们都是跨边界的像素，并对与包含这些像素的面积成比例的两个主要亮度级进行了平均，如图 7.30 所示。

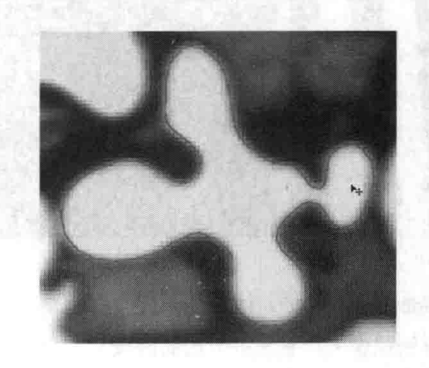


图 7.30 跨越边界线的有限像素示例，亮度值是根据每个区域内像素的面积来对两个区域中的亮度值平均得到的

直方图仅显示了不同值的出现频率；它不保留关于位置、相邻像素亮度、局部亮度梯度的振幅和其他因子的任何信息。确定边界位置仍需要空间信息。对于根据像素亮度间的距离来计算像素亮度的所有可能组合的图像，我们可为其建立一个共生矩阵。该信息用于选取那些作为特征一部分的像素，而不是简单地依赖于像素亮度值来选取像素。

使用直方图的形状来选择阈值时，通常会将阈值定在两个尖峰之间。问题是，由于直方图的这个区域非常低，几乎没有几个像素有这些值，因此计数统计结果很差，而且直方图中曲线的形状很

难定义。造成的结果是，阈值很难定位，而且它可能会随着光照的很小变化而移动相当大，视野中的变化包括具有不同形状的物体，或沿边界出现的一些像素。使用多项式拟合方式平滑直方图，一定程度上可更精确地找到最小点。

图 7.31 显示了一幅测试图像，图像中有两个视觉上可以分辨的区域。每个区域中均包含有标准差相同但均值不同的高斯噪声模式，且这两个区域中的亮度值重叠。这意味着将阈值设为两个尖峰之间的最小值，会使得两个区域中的某些像素会被错误地分类，如图所示。下面的几个例子中将使用这幅图像来对阈值处理和处理方法的效果进行比较。

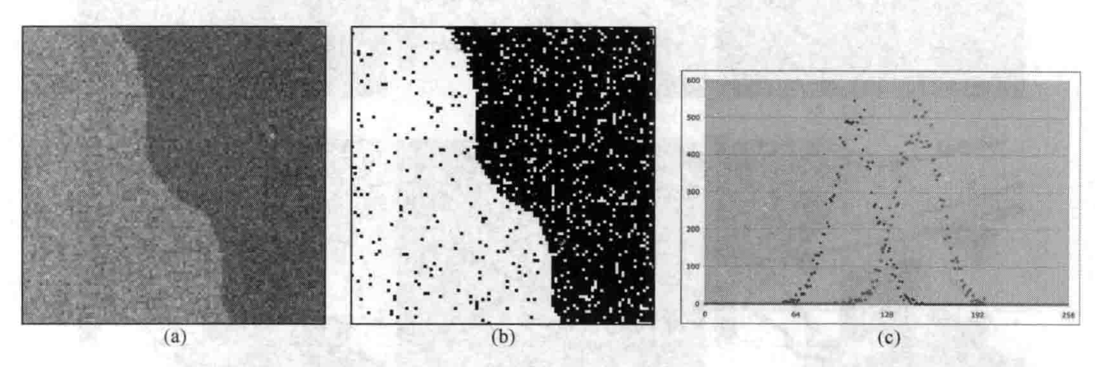


图 7.31　包含两个区域的测试图像，各区域的平均亮度不同，但在个别重叠的像素上有变化：(a)原始图像（为显示像素已放大）；(b)直方图，表明区域间有重叠；(c)以最小点处设置一个简单的阈值的结果

这类图像通常是在光子总数或其他信号很低的情形下得到的，而且计算统计在均匀区域中会导致像素亮度的变化，进而在直方图中会出现较宽的尖峰。计数统计会生成一个泊松分布，但当涉及相对较大的数量时，它就会非常接近于这些图像中使用的高斯函数。对于存在大量噪声的图像，如来自 SEM 的 X 射线点图，在试图进行阈值处理之前，可能需要一些其他的处理（O'Callaghan，1974）。在信号放大和数字化中，其他噪声源也会导致亮度发生变化，但形状上可能不是高斯的。

图 7.32 显示了一幅典型的稀疏 X 射线图。图像中的边界视觉上很明显，但它们的确切位置充其量是近似的，要求人类视觉系统将这些点都组合起来。成像处理可以通过统计围绕每个像素的一个圆形邻域内的点数，来确定边界。用一个直径为 5 像素、元素为 1 的核与图像卷积，可确定边界，结果如图 7.32(c)所示。对这幅灰度图像进行阈值处理，可找出所显示的边界，但没有足够的数据来判断边界中的那些小区域、空洞和不规则性是真实的，还是由有限的计数统计导致的。通常，阈值电平是通过求出背景区域中的平均亮度电平来设置的，然后设置标准差高于此阈值的几个阈值，来选取那些有意义的区域。图中比较了平均方法与中值滤波降噪声方法［见图 7.32(e)］，然后对图像进行了阈值处理，并应用了形态学编辑（见第 8 章）。两个过程的结果［见图 7.32(h)］略有不同，但原始图像中包含的数据确实不足以区分它们。

可以用几种方法来改善带噪区域的分割，图 7.31 就是这样的一个例子。第 8 章将讨论二值图像的编辑操作，包括形态学处理。首先对图 7.31(b)中的二值图像进行阈值处理，然后进行一系列的膨胀－腐蚀运算（称为闭运算），填充孔洞，删除孤立的像素，并平滑边界线，就产生了图 7.33(a)所示的结果。

为进行比较，应用到图 7.31(a)所示原始灰度图像的一种运算，基于两个区域的熵，将一个区域中的像素重新分配到了另一个区域中。此时，将像素集合到两个区域中可视为一个熵问题（Kanpur et al.，1985）：每个区域的总熵可计算为 $-\sum p_i \log p_i$，其中 p_i 是亮度为 i 的像素的百分比。求解将每个像素分类到两个群落之一中的边界，可为这两个区域最小化该函数，约束条件是每个区域中的像素必须彼此

接触，产生如图 7.33(b)所示的边界线。其他约束，如不同类中接触像素的数量最少，会平滑该边界。问题是，这种约束假设已知标本的一些内容（如已知边界是平滑的），且通常能以其他方式更有效地应用（例如通过平滑二值图像）。

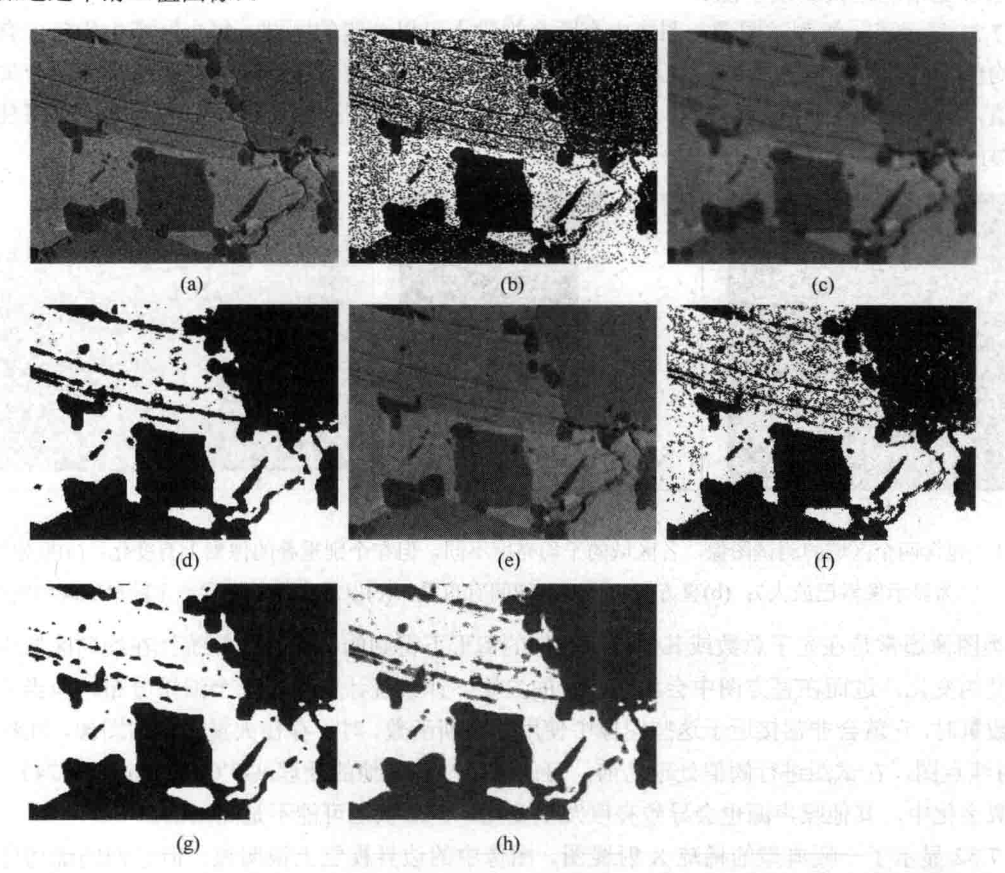


图 7.32　勾画出带噪图像中的区域：(a)原始 X 射线图（云母中的氧气）；(b)对图(a)应用阈值处理后的结果；(c)原始图像经 2 像素半径平均滤波器平滑后的结果；(d)图(c)经阈值处理后的结果；(e)原始图像经 2 像素半径混合中值滤波器降噪后的结果；(f)图(e)经阈值处理后的结果；(g)图(f)经形态学闭运算后的结果；(h)图(d)阈值处理结果与图(g)置入颜色通道后结果的比较

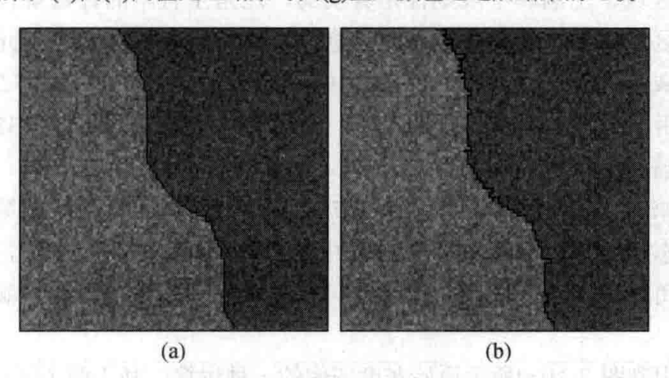


图 7.33　由各种方法确定的图 7.31 中图像的边界：(a)以直方图最低点进行阈值处理，然后闭运算（膨胀和腐蚀）的结果；(b)交互地设置最小熵点

在直方图中的最低值处设置一个阈值，有时称为选择最小面积灵敏度（Weszka, 1978; Wall et al., 1974）。

这意味着改变阈值时，不会明显地改变特征（或背景）的面积，但这种说法与添加到特征中的像素或从特征中删除的像素的空间排列无关。直方图中的任何最低点按定义是具有最小面积灵敏度的点。

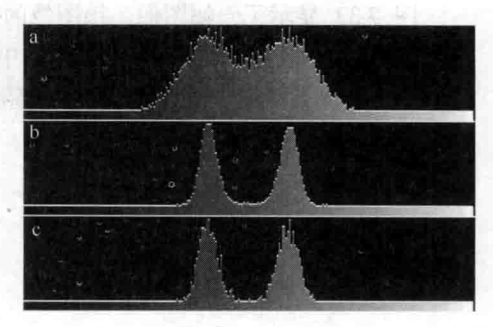


图 7.34 图 7.31(a)所示图像的直方图：(a)原始图像的直方图，两个尖峰间有重叠；(b)平滑后的直方图；(c)中值滤波后的直方图

对图 7.31(a)所示图像进行处理，可以改变直方图，产生更深、更宽且有一个更稳定最低值的最小值。图 7.34 显示了对图像平滑（用标准偏差为 1 个像素的高斯核）或应用半径为 1 像素的中值滤波器的结果（这两种处理均已在第 4 章中介绍过）。各尖峰更窄，谷更宽更深。图像的结果，以及通过在两个尖峰间设置阈值电平所选取的边界，如图 7.35 所示。将各边界线放到不同颜色通道中，会发现它们在位置上有 1 个或 2 个像素的不同，但原始图像中未提供足够的信息来让我们在它们之间选择。

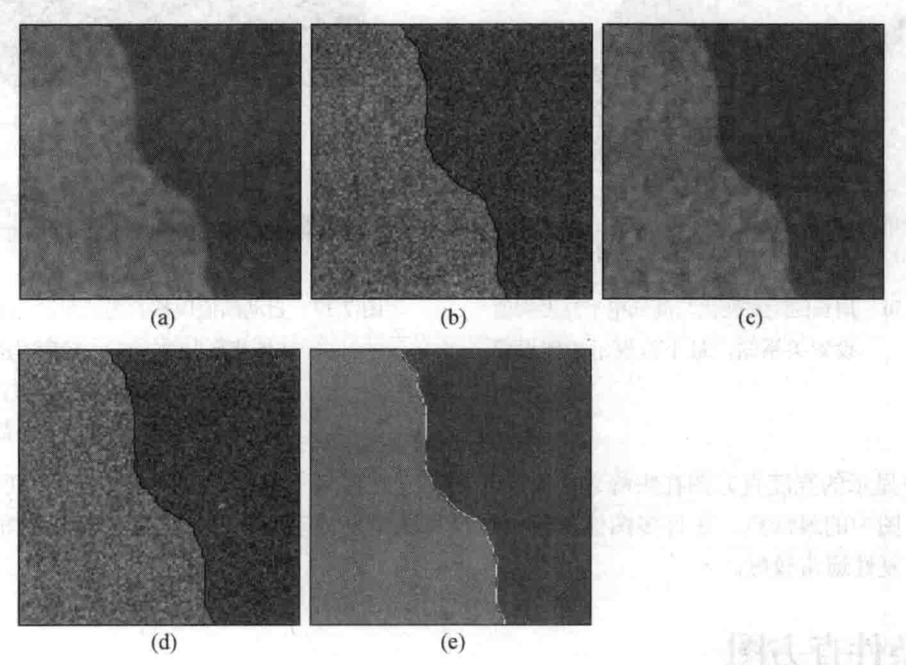


图 7.35 对图 7.31(a)所示图像进行阈值处理：(a)使用标准差为 1 像素的高斯核平滑图像后的结果；(b)图像(a)经阈值处理后生成的边界，已叠加到原始图像上；(c)中值处理（反复应用直到不再出现变化）结果；(d)图(c)经阈值处理后生成的边界，已叠加到原始图像上；(e)将这些边界线放到彩色通道中后，图 7.33(a)中的边界（红色）、图 7.35(b)中的边界（绿色）和图 7.35(d)中的边界（蓝色）的比较（白色、青色、黄色和品红色像素表示各个边界是重叠的）

图像中不同区域的面积（即直方图所表示的内容），看起来确实不能是熟练人工操作员观看一幅图像并交互式地调整阈值时所采用的标准。随调整变化很小的特征总面积，对于人工判断比较困难；另一种方法是使用特征周长的总变化（Russ & Russ，1988a），这对应于“最平滑的”边界。总周长随阈值的变化提供了一个可以有效计算的客观标准。这种响应曲线上的最小值提供了一种可重复设定阈值方式，以适应变化的光照等，并在一定程度上模仿人类假定边界应该平滑时来设置值的方式。对于上下阈值电平都要进行调整的情形，这会生成一个二维响应曲面（上下值），对该曲面求解，可找到最低点，如图 7.36 所示。

图 7.37 显示了一幅图像，该图像的亮度阈值已自动定位，以使总边界长度的变化最小。（蛋黄酱上的油滴）标本就是这样一幅图像，其中由于油－水界面的表面张力，我们期望边界是平滑的。许多情形下会出现类似的平滑边界，如生物组织的细胞膜、材料的相位变化（凝固或熔化），但这样的假设对于磨损或断裂的粗糙表面是不适合的。

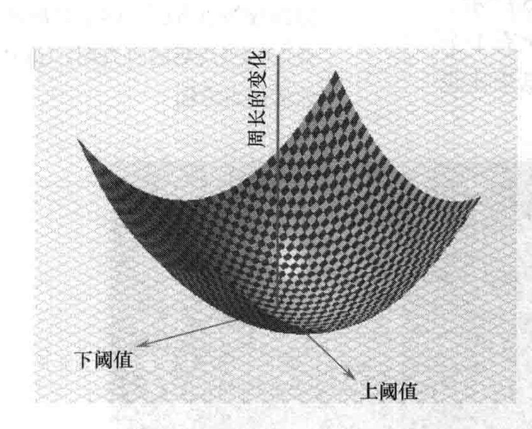


图 7.36　周长的绝对变化与高低电平亮度阈值设置关系图。最小值表示最佳设置

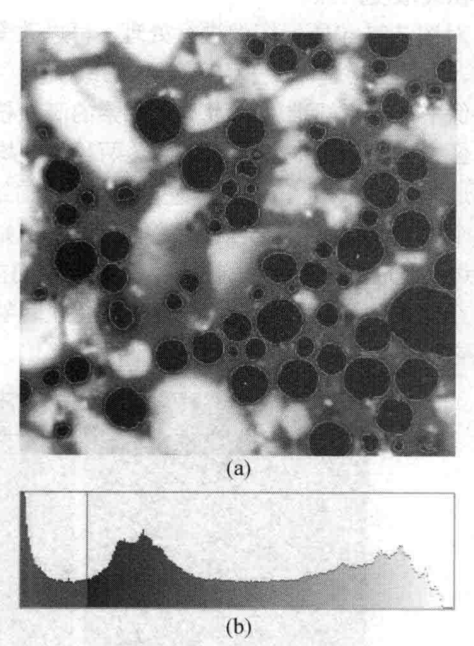


图 7.37　自动阈值调整的测试图像：(a)原始图像（蛋黄酱上的油滴），轮廓对应于图(b)中的阈值设置；(b)亮度直方图，显示了根据周长变化最小确定的最优阈值点

图形中显示的亮度直方图在尖峰之间有一个谷，它在形状上不是对称的或高斯的。所选的阈值点并不是直方图中的最低点。在许多图像上使用这个算法重复测量表明，在中度图像噪声和光照改变的情况下，重复性通常较好。

7.10　条件直方图

在用典型直方图表示的区域之间选择最佳亮度阈值时，大部分困难源于直方图的中间亮度值。这些像素中的许多沿两个区域的边界出现，因此从直方图中理想地消除它们的方法是，只保留来自均匀区域的尖峰，进而便于选取一个阈值（Weszka & Rosenfeld，1979；Milgram & Herman，1979）。

进行这种选择的一种方法是使用另一幅导数图像，如 Sobel 梯度或如第 5 章中介绍的任何其他边缘查找算子。有较高局部梯度值的像素，通常位于区域间的边界上，降低原始图像直方图中两个相位尖峰间的背景电平，可将这些像素消除。图 7.38 显示了一个例子。原始图像中包含有不同灰度级的三个相位。消除边缘像素的方法有几种。对梯度图像进行阈值处理，可选取具有较高值的像素。这会生成一幅二值图像，它可用做掩模，以将原始图像中的像素限制应用到待分析的直方图中。

例中，选择 Sobel 梯度图像中 20%具有最大幅度的像素来生成一个掩模，进而从原始图像和直方图中删除这些像素。结果是直方图中尖峰间的那些部分减少了，而尖峰本身几乎未受影响。这就使得由各个相位并在它们之间选择一个一致的点来表征尖峰的形状更为容易。

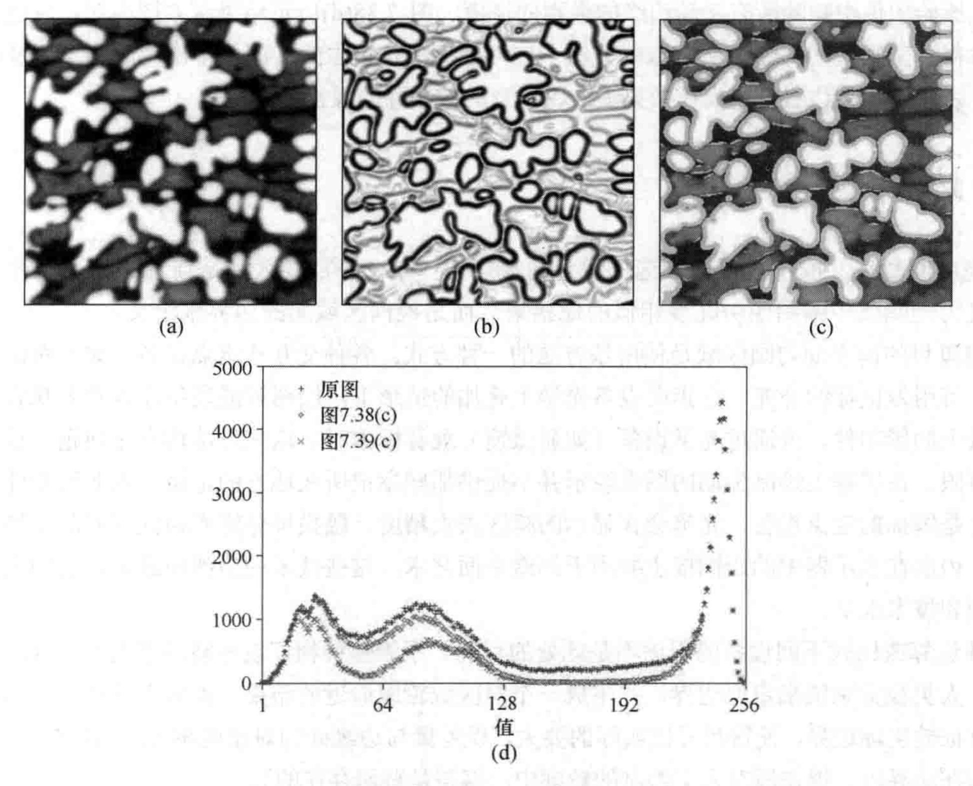


图 7.38　通过忽略边界像素进行阈值处理：(a)原始图像，它包含有三个具有不同平均灰度级的相位区域；(b)对图(a)应用 Sobel 梯度算子后的结果；(c)20%的像素具有最大梯度值的图像，它消除了原始图像中跨越边缘的像素；(d)图(a)中原始图像的直方图与图 7.39(c)中图像的直方图比较，主峰间的像素数已减少

这种方法需要在梯度图像上设置一个阈值来选择将要忽略的像素，方法通常是选择具有最高梯度值的某个固定百分比的像素。在所示例子中，梯度算子对白色和灰色区域之间较大差异的响应，要比灰色和黑色区域之间较小差异的响应，强烈得多。因此，跨越边缘的像素（及直方图中它们的背景）在白色和灰色尖峰之间比在灰色和黑色尖峰之间降低得更多。

图 7.39 显示了另一种方法。从图 7.38(b)所示的梯度图像开始，先使用非极大值抑制（又称灰度细化、骨架化或寻脊）使边界变窄，并消除实际上不在边界上的像素。这条线均匀地膨胀到 3 个像素宽，

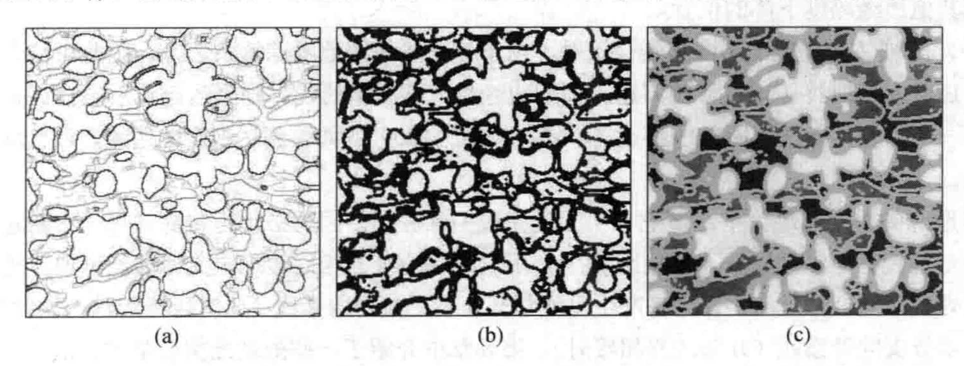


图 7.39　边缘像素的删除：(a)图 7.38(b)经非最大抑制（灰度细化）后的结果；(b)边界线膨胀到 3 像素宽；(c)保留原始图像中均匀内部区域，消除了边缘。经掩模处理后的图像的直方图，如图 7.38(d)所示

并用做从原始图像中删除所有跨越边缘的像素的掩模。图 7.38(d)所示结果直方图表明，灰色和黑色尖峰间的谷得到了很大的抑制。所有这些方法一定程度上都是自主的；图像中出现的不同区域亮度的特殊组合，表明了哪种边缘查找操作效果最佳，以及应被删除的像素。

7.11 边界线

按亮度或颜色，或仅按位置来选择像素的缺点之一是，不需要区域是连续的。另一种方法不将区域定义为一幅或多幅图像中亮度相似的像素集，而是根据区域间的边界来定义。

因测量目的而手动勾画区域是使用该方法的一种方式。各种交互式定点设备，如绘图板、触摸屏或光笔，可用做鼠标的补充，在指向设备光学上叠加的情形下，绘图可能发生在查看者观看计算机屏幕、平板上的影印件，或通过光学设备（如显微镜）观看标本时。这些方法均存在问题。显示器的分辨率很有限。在屏幕上绘制生动的图像表示并不提供那些标记所在地方的记录，除非该软件画的是直线。鼠标是笨拙的定点设备，光笔会在显示的暗区失去精度，触摸屏分辨率很低（同时手指会挡住视线）等。内嵌在显示器中的绘图板主要用于创造平面艺术。这些技术的最终问题是，它们依赖于观察者的判断和技术水平。

描述运算或比较不同设备的用途不是此处的目的。不管哪种物理设备被用于人工勾画，这种方法都依赖于人类视觉系统来定位边界，并生成一个对区域轮廓逼近的结果。许多人往往只画出他们认为重要的特征的实际边界，使得尺寸比实际的要大，误差量与边缘处的对比度相关（当然存在一些例外。有些人画在边界内。但在所有人工绘制的轮廓中，偏差是普遍存在的）。

人工勾画的其他问题包括由于期望而对偏差的敏感，不同人在不同时间所看到结果的重复性差，需要大量时间去标记出现的所有对象和结构。

使用计算机算法模仿人工勾画操作需要一个起点，这个起点通常是由人提供的。然后，该程序检查每个相邻像素，找到哪些像素具有边界的特征，并通常定义为亮度阶。选择具有最高局部梯度值的像素，将它添加到生长多边形中，然后重复该过程。有时会增加约束条件来最小化急剧变向，如根据方向来对像素值添加权重。

自动边缘跟踪也存在一些问题。首先，边缘定义本质上是局部的。人们具有相当强的适应能力来预判各种距离，进而找到将要连接到一起的边缘片断。完形心理学家称此为组合，详见第 2 章。对于只针对一个小邻域的算法，很难得出这样的结果。即使是在相当简单的图像中，沿边界也可能存在局部梯度或其他边缘测度下降的位置。

此外，区域邻接位置的边缘可能会接触。由于该算法等同地跟踪每条边缘，因此可能会给出荒谬的结果。还存在何时终止该过程的问题。如果边缘是一条简单的线，则到达起点后它会终止。如果该线到达另一个特征（上次应用程序已定义了该特征的边界），或如果到达视野的边缘，那么就无法完成这一勾画。

边缘跟踪的主要问题如下：(a)它自身无法完成图像分割，因为必须为它提供每一个新起点，而且它不能确定是否有更多的轮廓需要被跟踪；(b)用于跟踪边缘的相同边缘定义标准，可通过处理整幅图像和阈值处理而更容易地应用。这将产生可能断开和不完整的像素线（此时边缘跟踪无法继续），或产生具有许多分支的像素线（几条边界相接时）。第 8 章中介绍了一些处理此类缺陷的方法，这些方法使用的是腐蚀/膨胀形态学或分水岭分割技术。处理运算的全局应用会找到所有边界。

图 7.40 说明了几个这样的影响。该图像由几条手绘的暗线组成，对该图像添加了少量的噪声，并

应用了脊线跟踪算法（Van Helden，1994）。图中显示了每个用户选择的起始点以及由自动程序跟踪的路径。用于该例的设置，指导算法考虑向外 5 个像素距离的点，以决定在每个点处的移动方向。增加这一数字会产生人为平滑的边界，但需要更多的时间，因为此时必须搜索更大的邻域。反之，减小这一数字，会使得算法更有可能跟踪错误的拐角。很多路径是成功的，但也有很多路径并不成功。而对图像进行阈值处理来选择暗像素，然后对所得轮廓进行骨架化处理，如第 8 章中所述，一次就对所有的区域产生了合理的边界。

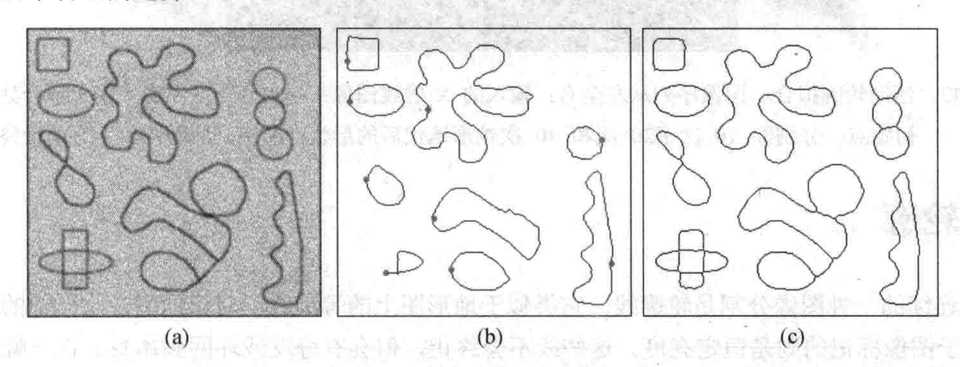


图 7.40　用于自动跟踪边界的测试图像：(a)添加了随机噪声的手绘线条图；(b)自动跟踪找到的线条，显示了每条线的起点（注意未跟踪到交叉线或分支线的某些部分）；(c)阈值处理和骨架化处理后找到的线条

对真实图像可以进行相同的比较。图 7.41 显示了来自光学显微镜的荧光图像。在这种情形下，自动脊线跟踪方法找到的不完整边界线，已通过手动辅助技术进行了补充。用户在边界附近画一条线，然后算法移动到最近的最暗点处。该方法［有时称为“主动轮廓”或“蛇行”(Kass et al.，1987；Lee，2006；Myronenko & Song，2009)］允许用户克服许多困难来使自动方法保持正确的方向。但使用阈值处理和骨架化方法得到边界线的速度更快，尽管所得边界线的细节不同，但哪种方法更优则无定论。

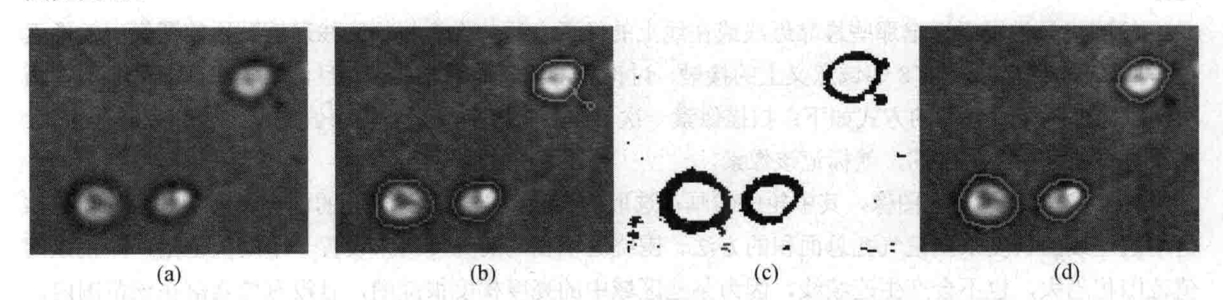


图 7.41　具有三个特征的光学显微镜荧光图像：(a)原始图像；(b)边缘跟踪算法圈出的轮廓叠加到原始图像上后的结果（红色为自动算法圈出的结果，绿色是人为圈出的特征）；(c)对原始图像进行亮度阈值处理的结果；(d)图(c)骨架化处理圈出的轮廓叠加到原始图像上后的结果

由于对边界上的所有点的拟合是使用最小化技术一次性完成的，因此蛇形技术可容纳一些丢失或混淆的点。蛇行技术尤其适合于跟踪系列图像中的移动边界，因为不同时刻的蛇形变化很小。任意一条边的最小长度、任何顶点处的最大角变化以及其他任意拟合常数，必须反映将被拟合区域的一些独立知识或判断。对每种应用，通常必须对它们进行调整，直到所得结果与视觉检查结果相比是可以接受的。图 7.42 显示了一个例子。这种自适应性通常使得蛇形技术适用于跟踪机器视觉中的已知移动对象。当应用到三维像素阵列（如三维医疗成像）时，蛇形形状会成为可变多面体，并通常称为气球。

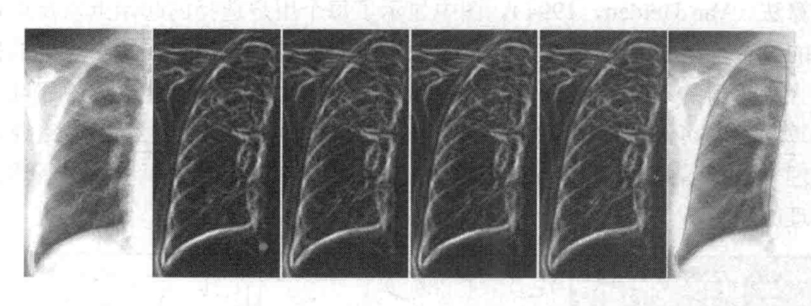


图 7.42　活动轮廓拟合。图像序列从左至右：输入的 X 射线图像；叠加到梯度图像上用于拟合处理的初始点；分别经 10 次、20 次和 40 次蛇形迭代后的形状；叠加到原始图像上后的最终结果

7.12　轮廓

保证连续的一种图像分割是轮廓线。它类似于地形图上的等高线。对于地图，它标记的是恒定高程，而对于图像标记的则是恒定亮度。这些线不会终止，但会有分支或环回到本身。在一幅连续图像中或实际的地形面上，总是存在轮廓线通过的一个点。对于离散图像，轮廓线的亮度值可能并不恰好对应于特定的像素值。尽管如此，如果存在一对像素，其中一个值亮于轮廓线电平，另一个值暗于轮廓线电平，那么轮廓线一定在它们之间的某处穿过。

原理上，轮廓线可拟合为通过这样一些点的多边形，这些点是在所有包含该轮廓值的像素对的中间内插出来的。这就允许我们以小于 1 像素的尺寸来测量这些线的位置，以及它们所表示边界的位置，此种方法称为子像素抽样或测量。第 10 章会使用这种方法来进行特征测量，但这种方法通常不用于表示整个幅，因为它需要由这样的一系列点来表示每条边界，而这些点必须集中成一个多边形。

轮廓线最常用于标记那些最靠近线或在线上的像素。这些像素将轮廓线逼近到原始图像中的像素分辨率，形成相互接触（8 邻域意义上的接触，讨论见下）像素的连续带，且可用于在许多情形下勾画出特征。由图像生成线的方式如下：扫描像素一次，然后将每个像素与轮廓值左上方的邻点进行比较，如果测试值在两个值中间，就标记该像素。

图 7.43 显示了一幅图像，其中轮廓线画在选取的亮度值处来标记气孔的边界。第 9 章中将介绍如何用这些线的长度来测量气孔总面积的方法。围绕这个相同的亮度级来设置一个阈值范围，即使该阈值范围相当大，也不会产生连续线，因为某些区域中的亮度梯度很陡峭，且没有像素落在该范围内。其他区域中的亮度梯度非常平缓，因此使用诸如 Canny 这样的边缘查找算子不会显示所有相同的边界，也不会引入更多的噪声。

在图像上画出一系列轮廓线，是显示图像中亮度较少变化的一种有效方式，如图 7.44 所示。甚至对于如图 7.45 所示的复杂三维场景，一系列轮廓线通常也能够勾画出相似的区域，或传递结构的意义。距离图像作为一类重要的图像，其中的像素亮度表示高程，这样的一组线就是地形图。此类图像由雷达或激光雷达成像、CSLM、干涉仪、STM 或 AFM 以及及其他设备生成。图 7.46 显示了硬币的一幅扫描探针图像，其中画出了轮廓线来描绘凸起的表面，另一幅类似的图像是圆珠笔尖图像。笔珠上的轮廓线显示了表面的粗糙度和圆度，可用于定量测量目的。第 15 章中将介绍此类距离图像的测量。

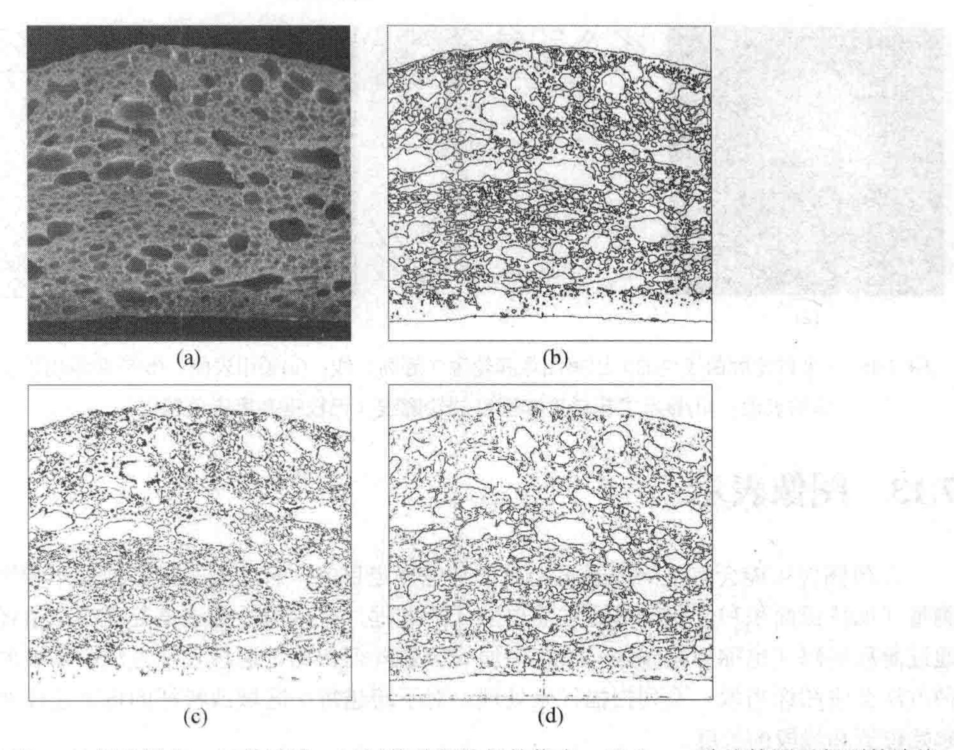


图 7.43　面包上的气孔：(a)原始图像；(b)轮廓线；(c)阈值处理选择的像素；(d)Canny 边缘检测子得到的边缘

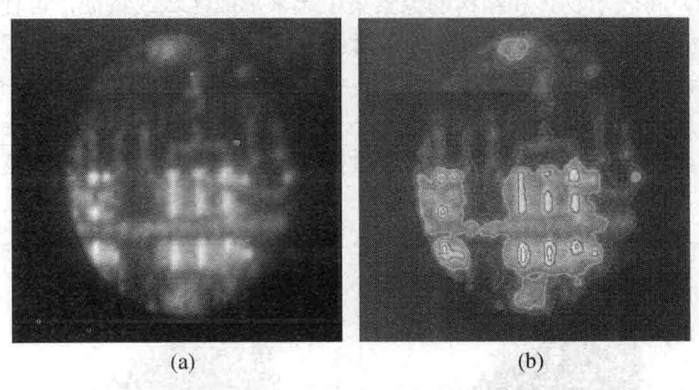


图 7.44　掺杂到晶圆中的硼的离子探针图像：(a)原始图像，其中的亮度与浓度成正比；(b)叠加等亮度或等浓度轮廓线后的图像，因此更易于对图像不同部分的值进行比较

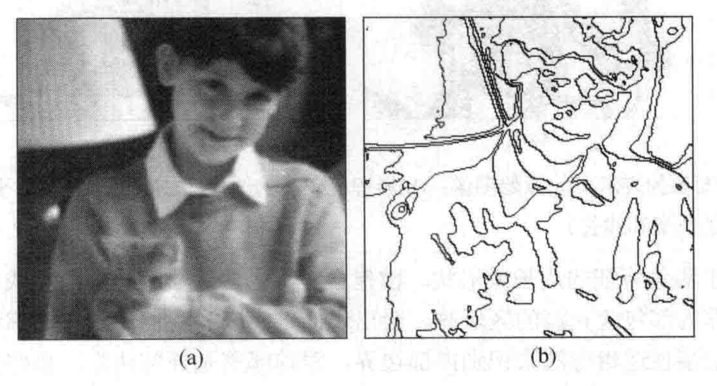


图 7.45　(a)真实图像；(b)在 4 个不同亮度值处画出的轮廓线。无论轮廓线变得多么不规则，都会是连续的和可区分的

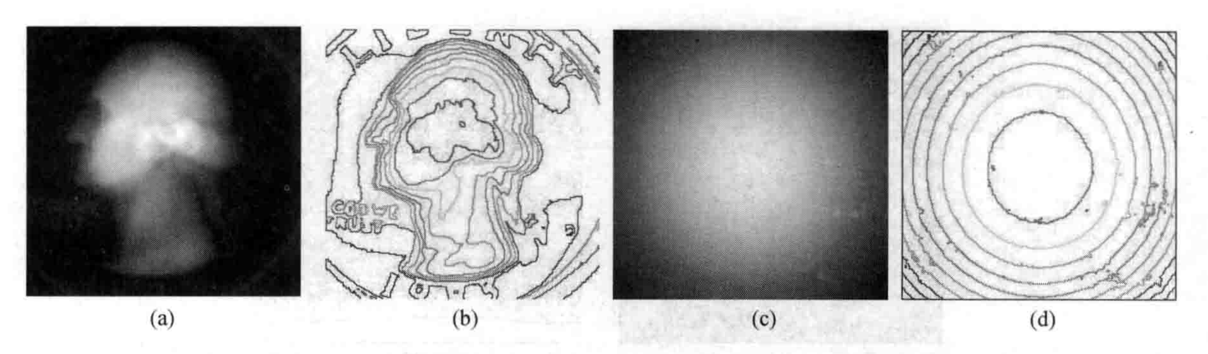
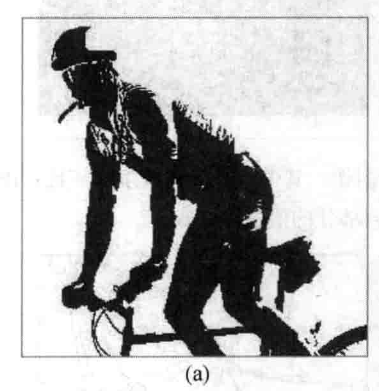

(a) (b) (c) (d)

图 7.46 （来自轮廓曲线仪的）距离图像和轮廓（等高）线：(a)硬币表面；(b)轮廓线描述了凸起的表面；(c)滚珠的表面；(d)显示了粗糙度和圆度的轮廓线（已根据高度彩色编码）

7.13 图像表示

二值图像可能会有不同的表示；对于某些特定目的，有些表示与其他表示相比会更有用。大多数测量（如特征面积和位置）可通过简单的计数过程，直接由经过阈值处理的二值像素表示来计算。与通过游程编码（也称为弦编码）来存储原始像素阵列相比，这种表示方法所需要的存储空间更少。这种方法会将图像当做一系列扫描行来处理。对于跨越每个区域或特征的每条连续线段，它存储线号、起始位置和线段的长度。

对于典型的图像，将表示物体和结构的像素集中起来形成区域或特征，可使得游程编码表远小于原始图像。例如，这就是通过电话线来传输传真消息所用的方法。图 7.47 显示了为此目的对一幅黑白图像编码的方法。在该例中，原始图像有 256×256 = 65536 个像素，但游程表只有 1460 个字节长。游程表可以直接用于测量面积和位置，其计算量要小于像素阵列的计算量。由于各弦的顺序对应于光栅经过特征的顺序，因此在出现几个特征时，需要按逻辑顺序来识别每个特征的弦，但这通常在表建立后进行。


(a)

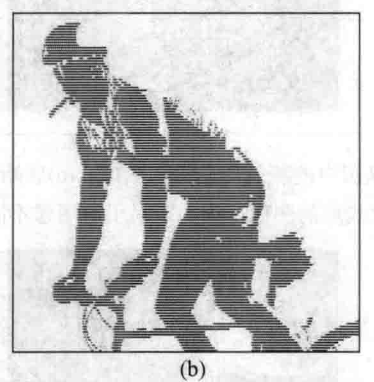
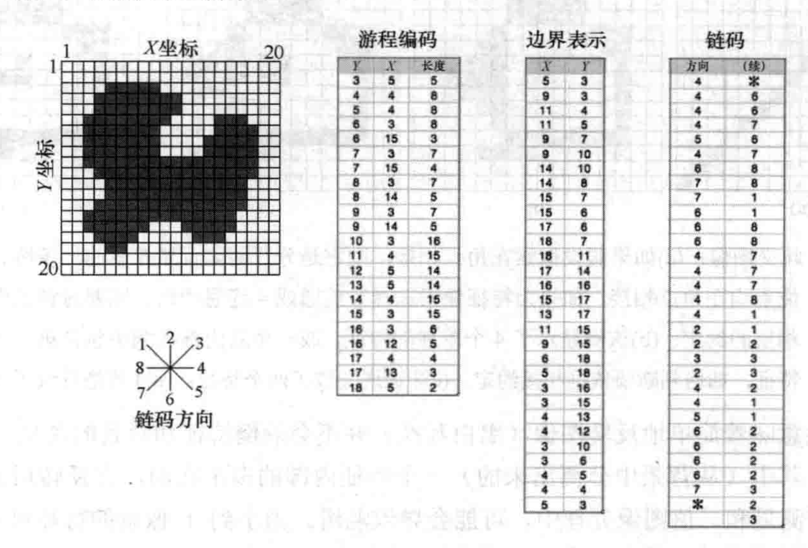

(b)

图 7.47 黑白传真图像的表示：(a)原始图像；(b)游程编码后的图像（红点表示每条水平线的起点，只发送每个红点的位置和线长）

弦表不太适合于测量特征的周长或形状。由包含边界的多边形的坐标所组成的边界表示，更胜任这一任务。但在处理内部包含孔洞的区域时，会出现问题，因为我们无法将内部边界与外部边界联系起来。因此必须使用某种逻辑方法来识别内部边界，跟踪哪些是外部边界，哪些是内部边界，并在需要时建立特征内各特征的层次关系。

需要时（如需要计算周长时），使用一系列弦的端点，可以由游程表生成多边形形式的边界。边界

多边形也可以由所有边界点形成，这些边界点由从一个边界点指向下一个边界点的一组短向量（链码）组成。在方形像素阵列上，每条这样的线的长度只能是 1 或 $\sqrt{2}$ 个像素，且只能是 8 个方向之一。对每个方向分配一个从 1 到 8（特定情况下也可从 0 到 7 或从−3 到+4）的数字，并按顺序为闭合边界写出所有数字来产生链码，如图 7.48 所示。



游程编码

Y	X	长度
3	5	5
4	4	8
5	4	8
6	3	8
6	15	3
7	3	7
7	15	4
8	3	7
8	14	5
9	3	7
9	14	5
10	3	16
11	4	15
12	5	14
13	5	14
14	4	15
15	3	15
16	3	6
16	12	5
17	4	4
17	13	3
18	5	2

边界表示

X	Y
5	3
9	3
11	4
11	5
9	7
9	10
14	10
14	8
15	7
15	6
17	6
18	7
18	11
17	14
16	16
15	17
13	17
11	15
9	15
6	18
5	18
3	16
3	15
4	13
4	12
3	10
3	6
4	5
4	4
5	3

链码

方向	（续）
4	*
4	6
4	6
4	7
5	6
4	7
6	8
7	8
7	1
6	1
6	8
6	8
4	7
4	7
4	7
4	8
4	1
2	1
2	2
3	3
2	3
4	2
4	1
5	1
6	2
6	2
6	2
6	2
7	3
*	2
	3

图 7.48　按游程长度编码、边界多边形表示或链码，来对二值图像中的相同区域编码

这种形式非常适合用于计算周长或描述形状（Freeman，1961，1974；Cederberg，1979）。统计偶数和奇数的数量，将奇数的个数乘以 $\sqrt{2}$ 来校正对角方向，然后相加，可求出周长。链码也包含形状信息，这些信息可用来定位角点、简化轮廓的形状、匹配与方向无关的特征，或者计算不同形状的描述子，详见第 11 章。

当前的多数成像系统使用方形像素阵列，因为方形像素阵列即适合于光栅扫描获取设备，也适用于处理图像并执行测量。如果使用一种沿扫描线而非线之间的不同像素间距获取了矩形像素，那么空间域中的操作（如邻域处理）或频率域中的操作都会更为困难，因为我们必须考虑到不同像素距离与方向间的关系。使用矩形像素也会使得测量更复杂。

使用方形像素阵列处理图像时，也会存在前几章中提及的一个问题：与 4 个直接相邻的像素相比，4 个对角相邻的像素到中心像素的距离要远一些。有些系统中使用的是另一种排列方式，即将像素放置为六角形阵列，其优势在于所有相邻像素间的距离都相等，同时共边的相邻像素间的距离也相等，这就简化了处理和计算。但其最大的不足之处在于，标准相机和其他获取与显示设备均不支持这种工作方式。

对于传统的方形像素阵列，必须确定角点处的相邻像素实际上是否相接。这对于第 8 章中的二值处理运算是非常重要的。如上所述，要将像素连接成特征或跟踪边界周围的点，这是有必要的。尽管不同阵列各有优缺点，但不管使用何种阵列来表示构成特征的像素，都须使得背景（围绕特征的像素）与特征的特性相反。

图 7.49(a)显示了这种双重情况。如果假设特征中的像素与其 8 个相邻像素中的任意一个相接（称为 8 连通），那么图形中的那些像素会分开两侧的背景，而对角相邻的背景像素则不相接。因此，它们是 4 连通的（只与它们的 4 个共边邻点相接）。相反，如果背景像素对角相接，那么这些像素是孤立的，且仅沿它们的边缘相接。对于图 7.49(b)中的灰色像素，选择 8 连通规则，生成了两个特征，其中一个特征内有一个孔洞，而第二个特征则在这个孔洞内［见图 7.49(c)］。若使用 4 连通规则［见图 7.49(d)］，

就会生成 4 个特征，且现在 8 连通的白色背景是连续的。如下章节中给出的多数例子中，会对特征使用 8 连通规则，而对背景则使用 4 连通规则。

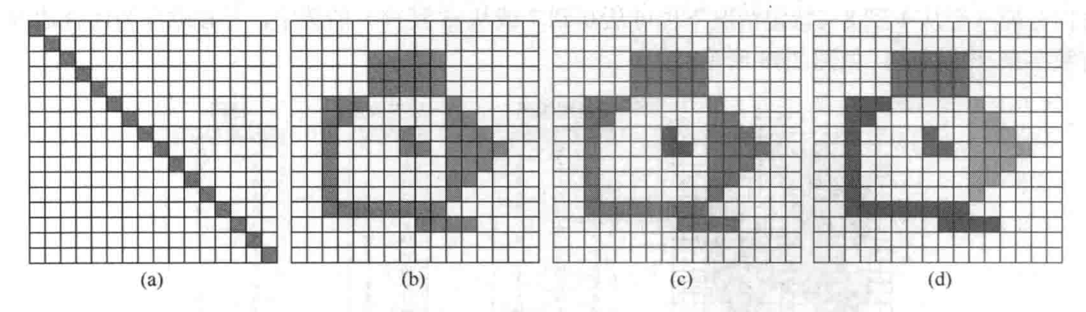


图 7.49　歧义图像：(a)如果假设像素在角点相接，则它是分开两侧背景像素的一条线。但这些像素均在角点相接。如果为特征像素选择 8 连通或 4 连通规则，则要对背景像素应用相反的约定；(b)该例显示了 4 个单独的特征，或一个其内部孔洞中包含第二个特征的特征，如何判断要依据相接约定；(c)8 连通导致了两个特征；(d)4 连通导致了 4 个特征

这种双重性意味着简单地反转图像（黑白互换）并不会颠倒特征和背景的含义。图 7.50 显示了这样一种情形，其中（从背景中分离出来的）一个特征内部的多个孔洞，在反转后的图像中变成了单个区域。这在测量和二值图像处理中，可能会导致混淆。当小到 1 像素的特征尺寸都很重要时，就会出现基本的不确定性。使用小像素的大阵列来精确定义小尺寸并表示特征拓扑时，这不可避免地会存在争议。

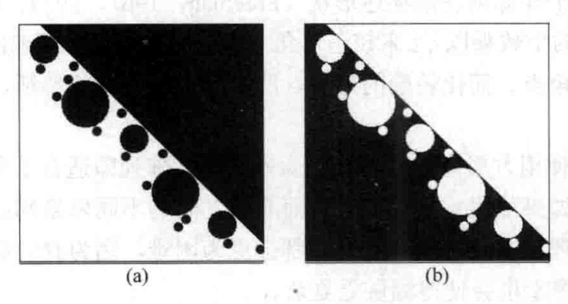


图 7.50　在不改变连通规则的情况下，反转图像会改变含义。在图(a)中，前景黑色像素都在角点相接（8 连通），所以这是一个具有不规则边界的特征。在图(b)中，背景白色像素并不相接（4 连通），因此它们是该特征内的多个分开的孔洞

7.14　其他分割方法

除上面给出的基于阈值处理的方法外，分割图像还有其他方法。这些方法通常会在机器视觉和机器人的意义上来理解图像（Ballard & Brown，1982；Wilson & Spann，1988）。两种最广泛使用的方法是分裂－合并和区域生长，尽管它们在方法上看起来正好相反，但其实具有许多共性。

分裂－合并是一种从整幅图像开始的自顶向下方法。首先选取一些图像性质，作为确定内容是否统一的标准。该标准通常基于亮度直方图的统计，但也可能会用到颜色信息或局部统计性质。如果直方图是多峰的，或具有较高的标准差，那么假设区域（初始时为整幅图像）是不一致的，并将该区域分成 4 个象限。需要时，可以相同的方式来审视每个象限并继续细分，直到覆盖了各个像素。父区域和其 4 个象限或子区域的关系，通常以四叉树结构来编码，四叉树是这种方法的另一名称。

这不是划分父区域并对所得的数据结构编码的唯一方法。可以用阈值处理技术来将每个区域划分成任意多个子区域，子区域又可以继续细分。这样生成的最终结果中，就很少有块状边界存在，但数据结构相当复杂，因为必须定义所有的区域，且所需的处理时间会非常长。

单独细分区域并不能创建有用的图像分割结果。在每次细分后，将每个区域与更高层级中不同方形中的相邻区域（通常只是共边区域）进行比较，如果它们相似，则将它们合并在一起。“相似”的定义可以使用应用于分裂运算的同样测试，或者只对沿着或接近共同边界上的像素进行比较。后者的优点在于可容纳跨越图像的渐变。

图 7.51(b)显示了一个例子，其中只进行了 4 次迭代。一些大区域已被合并，而且随着迭代的进行，其边缘会越来越清晰［见图 7.51(c)］。图像的其他部分包含了在区域变得可见之前需要进一步划分的各个方块。

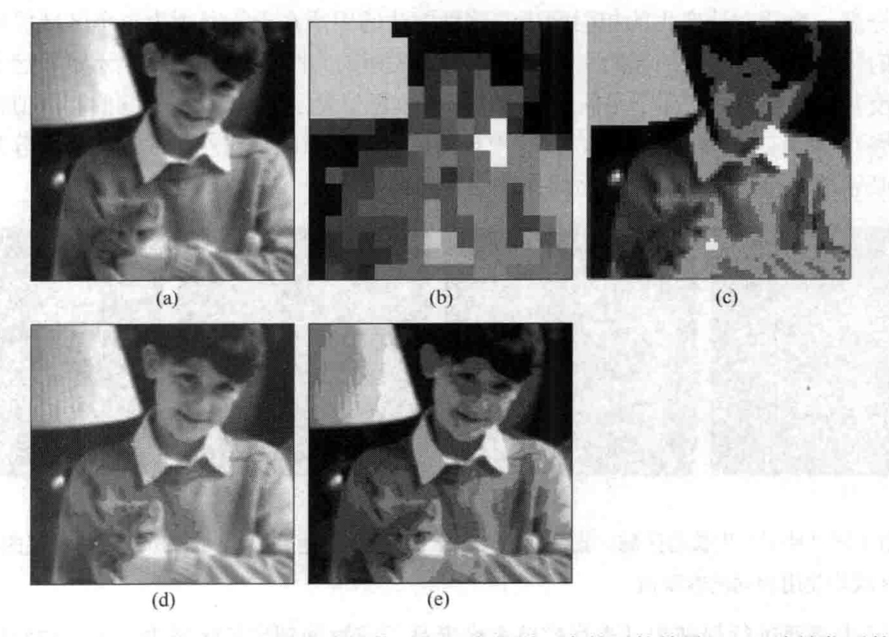


图 7.51　其他分割方法：(a)原始灰度图像；(b)4 次分裂—合并后的结果；(c)6 次迭代后的结果；(d)由女孩毛衣中的一个点生长的区域（以红色圈出）；(e)选取多个区域后的结果

分裂—合并这种方法的一个优点是，经过有限次数的迭代之后可实现完整分割（例如，一幅 512 像素的方形图像只需 9 次迭代就能到达单个像素，因为 $2^9 = 512$）。同时，区域和子区域的四叉树表可以用于某些测量目的，而且分割可一次辨识所有不同类型的区域。通过比较，阈值处理方法一次通常只能分离一种类型的区域或特征。对于包含多类物体的图像，需要多次应用该算法才能分离出区域或特征。

另一方面，分裂—合并方法取决于用于检测每个区域的不均匀性的测试的质量。这种方法容易忽略大均匀区域中的小子区域。例如，假定区域内部像素亮度的一种正态分布的标准统计测试几乎不适合于真实图像，因此必须使用更加复杂的过程（Yakimovsky，1976）。任何用于划分和合并区域的测试都可以表示为图像处理运算。经处理后的图像可以显示出用于分裂—合并测试的边缘与纹理，这些边缘和纹理与直接由阈值处理的方式得到的相同。但这种方法的效率低下，因为耗时的计算可应用到图像中的均匀部分，但结果是相同的。阈值处理还具有识别视野内不同部分中的类似物体的优点，而分分裂—合并方法则不能做到这一点。

相反，区域生长是一种从单个像素开始的自下向上方法。这种方法从一些“种子”位置（通常由

处理者提供，但有时由图像处理算法定位）开始，一次检查一个相邻的像素，如果它们足够相似，则将其添加到生长的区域中。再次对整个区域或局部像素进行比较，会发现后一方法允许亮度的平缓变化。在添加另一个像素的容限取决于生长区域中像素的标准差的意义上，或在支持最靠近正被计算像素的那些像素的加权平均的意义上，使得测试“自适应”是可能的，但会相当耗时。该过程一直继续，直到无像素可以添加为止。接着在另一个位置开始一个新的区域。图 7.51(d)显示了一个例子，其中已经识别出了一个区域；注意，它包括猫和小女孩的毛衣的一部分。图 7.51(e)显示了已识别出多个区域后的结果。

对于某个阈值范围内的区域，生长算法逻辑上等同于一个样本点的膨胀（尽管实现起来通常不同）。膨胀是一种形态学处理，详见第 8 章。

如果使用相同的比较测试来决定一个像素是否属于某个区域，那么区域生长过程的结果，会是与自顶向下的分裂－合并方法产生的相同区域。这种方法的困难在于，必须为每个区域提供起始点。取决于所使用的比较测试，不同的起始点可能不会生长成相同的区域。图 7.52 显示了看起来具有相同结构（披萨上没有烤的芝士）的几个不同起始点（以蓝色标记显示），以及用相同的容限值得到的不同区域（红色轮廓）。此外，在分类完成前，没有理想的结构来对由该过程产生的超出整个阵列的数据进行编码；由于必须逐个检查每个像素，因此完整的分类很慢。

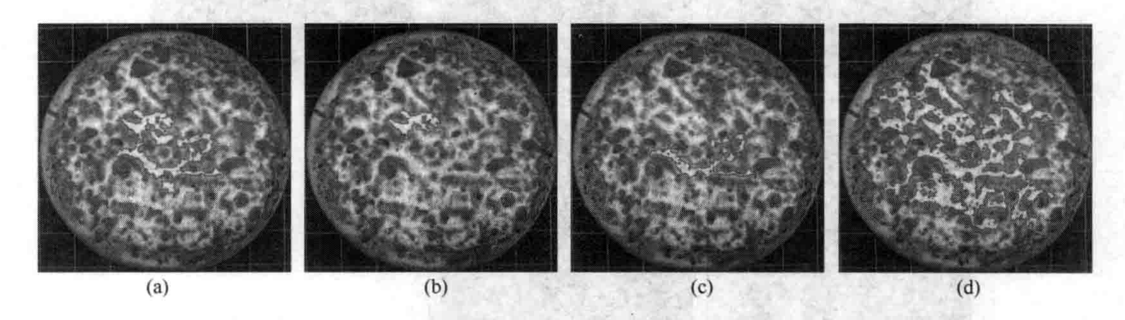


图 7.52　由不同“种子”生长的区域。蓝色标记表示原始种子，红色线条显示了所选连续区域的边界，所有区域均使用相同的容限值

区域生长也需要进行局部测试来观察单个像素是否应被加到生长区域中，并需要将测试放大到全局，以便能识别一些一致且明显的区域。如果使用的测试区域太小，那么常见的结果是，这些区域会跑到相邻的位置，或与不同的区域合并。甚至边界上的单个像素也能导致这种跑偏或合并。

最后要说明的是，不存在简单的方法来确定何时应完成该过程，以及确定已找出了图像中的所有有意义的区域。尽管与手动跟踪或边缘跟踪相比，区域生长（也称为“种子填充”技术）是选取图像中一些区域的有用方法，但它不适用于处理包含许多区域的复杂图像（Zucker，1976）。

7.15　一般分类问题

所描述的各种方法，在一定程度上依赖人类判断来识别区域的存在，并通过勾画边界或选择某个亮度值范围来定义它们。还有一些方法从不完整的定义开始，不断地细化分割，以达到更高的精度和一致性，完全自动的技术会确定有多少类物体出现，并将图像完全细分，以将这些物体隔离出来。但它们在小型计算机系统中应用甚微，且效率通常要比人工输入的效率低。普通图像分割任务可视为分类问题的一个例子。就像多数涉及要素或人工智能的技术那样，它可能不会使用与人类采用的相同输入或决策方法，但会力求实现相同的结果（且有时是成功的）。

卫星成像中一直使用了一种成功的一般分类方法，卫星图像中存在许多波段的数据（Reeves，1975）。如果在高维空间中画出图像中的每个像素点，空间中的各个坐标轴表示各波段中的亮度，那么对应于土地利用、作物类型、土壤或岩性等的不同类型的点将会聚集在一起，且这些聚类彼此能很好地分开，如图 7.53 所示。接下来，问题简化成寻找这些聚类，并拟合它们之间可以用来分类的边界。第 12 章中将详细探讨如何找到此类聚类和边界。

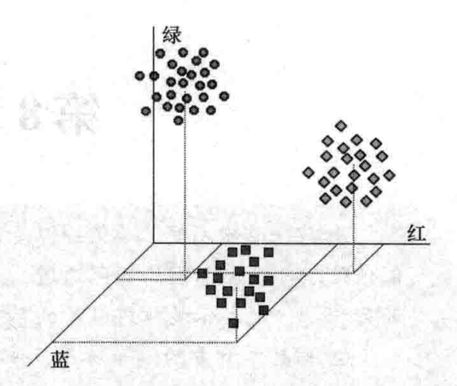


图 7.53 彩色空间中像素分类示意图。每个像素根据其颜色值在坐标中画出，聚类识别所出现的各个区域

简化到一维情形（一幅简单的灰度图像），该分类从亮度直方图开始。聚类分析找到尖峰，并试图在它们之间画出阈值。在有些特殊任务中，这是成功的，如统计显微镜载玻片上某类细胞的数量。随着维数的增加，例如使用来自彩色图像的 RGB 或 HSI 数据，或添加来自导数纹理或梯度图像的值，聚类的分离通常变得更加明显，但“尖峰”的位置及其边界的定义变得更困难。因此，该方法特别适合于具有几个可见和红外波段的卫星图像。

当聚类中包含许多类似的点时，它们更容易识别，但可能会忽视小区域或不常见的物体。此外，围绕聚类的背景点（在 n 维图中通常沿聚类之间的线分布）数量会对自动算法带来混淆。这些点是由跨越区域间边界的像素的有限大小产生的。找到几个主要的聚类可能很简单，但要找出所有聚类则很困难。

识别出聚类后（具有一些先验知识或人类输入会有很大的帮助），使用这些信息来对新点分类还存在不同的策略。一种是使用边界围起每个聚类，典型的边界是多面体或 n 维椭球。落入这些任意区域内的点立即就被分类。

对于椭球情形，也可使用一系列同心边界来包围聚类中不同数量的点，它可用于给出分类为新点的概率。这有时称为“模糊”分类方法。如果添加到聚类中的新像素改变了均值和界限，就会成为一种迭代方法：

第三种方法是，为每个新点找到最近分类的点，并对这个新点指派相同的聚类。这种方法有几个缺点，特别是存在分布稠密的聚类和成员很少的聚类时，或者聚类相互靠近甚至重叠时。它还需要很多时间来从大量的当前点中寻找并定位最近的那个点。该技术也有一种拓展应用，即识别少量最近的邻点，并“投票”来识别新点。甚至可根据现在值到待测像素的距离的倒数，来对它们进行加权（Chen et al.，2005）

为测量或识别目的将灰度图像分割成几个区域，可能是图像分析的最重要领域。正在使用的许多新技术是相当自主的，且适用范围有限。Fu and Mui（1981）和 Haralick and Shapiro（1988）中给出了大量此类文献。多数标准的图像分析教材，如 Rosenfeld and Kak（1982）、Castleman（1979）、Gonzalez and Wintz（1987）、Russ（1990b）、Pratt（1991）、Parker（1997）和 Russ and Russ（2007），都包含了关于阈值处理的章节。

在其他人工智能领域广泛使用了所有这些不同的方法及其改进（Fukunaga，1990）。它们可以在硬件、软件中实现，也可以在硬件和软件的组合中实现。目前，分割仍只用在少数人工智能技术中。但随着更多多波段成像的实现以及计算能力的增强，此类方法的使用会越来越多。

第 8 章　二值图像处理

如前一章中所述，二值图像是由根据某些性质选取的像素和像素组构成的。对亮度值进行阈值处理，组合不同的颜色通道，或在处理后提取纹理或其他信息，可完成这种选取。二值化的目标是，把表示物体或结构的特征从背景中分离出来，以便进行计数、测量或匹配运算。

但如第 7 章中的例子所示，分割运算的结果差强人意。对于复杂的真实图像，即使最好的分割程序也会把一些像素误分类为前景或背景。它们可以是沿区域边界分布的像素，或区域内噪声的结果，或恰好是具有所选特定性质或用于阈值处理的特性的像素。处理二值图像来校正这些阈值处理错误的主要工具有两类：合并图像所用的布尔运算，修改图像内各像素的形态学运算。

8.1　布尔运算

在第 7 章中关于彩色图像阈值处理的章节中，引入布尔“与”运算合并来自各幅彩色通道图像的数据。通过对每个 RGB 通道中的亮度值设置阈值，可让待选像素落在这些范围内。这种技术会生成三幅二值图像（每个通道一幅），这三幅图像然后由逻辑“与”运算来合并。这种处理过程会逐像素检查这三幅图像，为所选区域保留在所有三幅图像中都被选取的像素。

彩色阈值处理指的是这样一种情形，即比较和合并几幅不同图像（颜色通道）中同一位置的像素亮度值。在其他情形下，比较两幅或多幅图像中像素的位置和亮度值也是有用的。图 8.1 显示了一个例

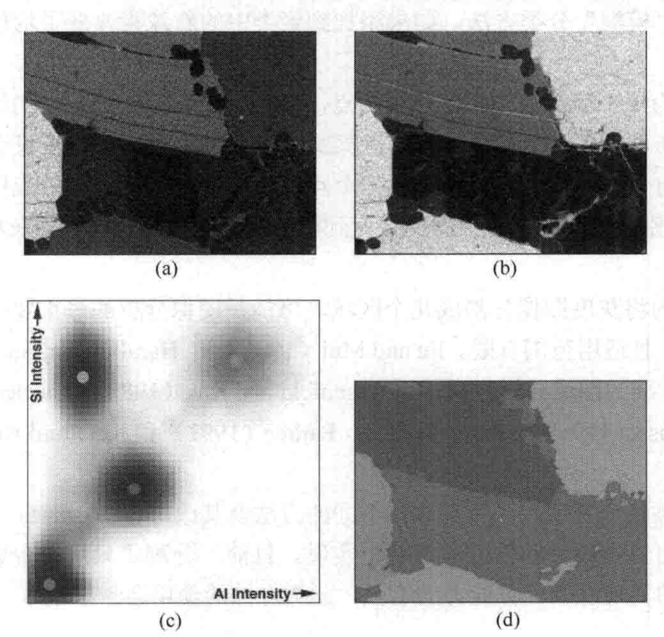


图 8.1　共位：(a)显示铝在矿物中的强度分布的 X 射线图；(b)硅的类似图；(c)共位图，其中每个单元中的像素数由灰度值表示，使用图(d)中的相应颜色标出了各个聚类；(d)图(c)中 4 个聚类中的像素位置

子。一块矿石样品上同一区域的两幅 X 射线图显示了强度分布，进而表示了铝和硅的集中分布。共位图（或共生图）使用两幅图像中每个位置的像素亮度值作为坐标来产生二维直方图显示，该显示中每个点的暗度表示具有那种值的组合的像素的数目。在结果图形的暗色区域中，有许多像素具有原始样本中元素集中度的相似组合。图中每个区域的大小，既受信号中存在的噪声影响（Adler et al., 2008），也受不同浓度中的任何真实变化影响。示例图中，基于硅/铝组合显示了 4 个区域，在原始图像中也可以观察到它们。

共位也用于多种染色制备的生物标本中，或用于以多个波长的光成像的生物样本中（Mossberg & Ericsson，1990；Manders et al., 1993；Zinchuk & Zinchuk，2006；Zinchuk et al., 2007；Zinchuk & Zinchuk，2008）。第 5 章中曾将它用做主成分分析例子的一部分。当一幅共位图显示具有同一位置亮度值的特定组合时，可对每幅图像进行阈值处理，并用“与”运算来合并两幅二值图像，进而生成所选区域的一幅图像。第 7 章中的图 7.22 和图 7.23 说明了多标准的“与”合并。

注意，此处所用的术语有前景、选中像素或“开”像素、背景、未选中像素或“闭”像素。选中像素是以白色、黑色还是以其他颜色显示，不存在统一的标准。许多情形下，在屏幕上将所选区域显示为黑色背景上的白色的系统，事实上可反相显示或打印为白色背景上的黑色。这种反相源于如下这个事实，即前景像素的选择与显示器或打印输出的某些操作相关。如果对图像有一定的了解，那么这种不一致性似乎不会给大多数用户造成困难。由于这里所用的许多图像不是常见物体，而且有些图像是示意性例子，因此在各种情形下连贯地定义前景像素（感兴趣像素）就很重要。这里用到的约定是，“开”像素或选中像素（它们构成用于后续测量目的的某个特征、物体或结构）显示为黑色，而“闭”像素（背景）显示为白色。

组合来自几幅图像或通道的信息时，“与”运算要求每个通道中位置 i, j 处的像素为“开”像素并显示在结果中。蓝色数量正确而红色数量不正确的像素被忽略，反之亦然。这就在二维空间中定义了一个矩形，或在高维空间中定义了一个棱柱体，其中含有将被包含的像素值。如前一章所述，在 n 维空间中为像素选取勾画出一个不规则区域，可描述更为复杂的颜色值组合。对离散范围进行“与”运算的优势在于，它使用二值图像可非常高效而快速地执行。

也可采用其他布尔逻辑规则来组合二值图像。4 个主要的逻辑操作是“与”（AND）、“或”（OR）、“异或”（XOR）和“非”（NOT）。图 8.2 中给出了这些基本运算。图 8.3 中给出了一些可能的组合。所

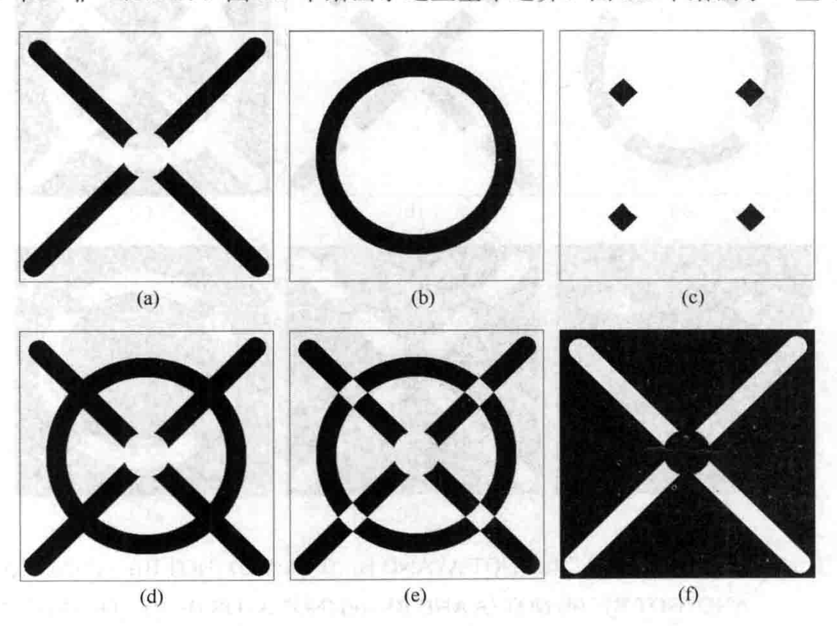


图 8.2　简单的布尔运算：(a)二值图像 A；(b)二值图像 B；(c) A AND B；(d)A OR B；(e)A XOR B；(f)NOT A

有这些操作都是逐像素执行的。由于涉及多幅图像的任何逻辑运算均可以分解为一次只使用两幅图像的一系列步骤，因此这里的说明只针对两幅图像的组合。图形中的说明类似于逻辑和集合论教学中使用的 Venn 图（也称为欧拉图）。

如上所述，“与”运算要求两幅原始图像中的像素都为“开”时，结果中才能为“开”。只在其中一幅图像中为“开”的像素，结果中将为“闭”。如果像素在任意一幅原始图像中都是“开”的，则“或”算子会使得该像素为“开”。

如果一个像素在任意一幅原始图像中为“开”，那么在 XOR 运算的结果中它为“开”；如果一个像素在两幅图像中都为“开”，那么在 XOR 运算的结果中它为“闭”。这意味着两幅图像的 AND 运算结果与两幅图像的 XOR 运算结果的 OR 运算，所得结果与先用 OR 运算的结果相同。实际上，产生相同结果的 4 种布尔算子的不同组合方式有很多。

AND、OR 和 XOR 运算需要两幅原始图像，并生成一幅图像作为结果。它们满足交换律，即两幅图像的顺序并不重要。A AND B 和 B AND A 的结果相同。NOT 算子只需要一幅图像。它只是简单地反转每个像素，即把“开”像素变成“闭”像素，或把“闭”像素变成“开”像素。有些系统实现 NOT 运算的方式是为每个像素交换黑值和白值。对于处理像素级细节，这可以正确地进行。对于特征级组合，如下所述，则必须考虑（第 7 章中讨论的）一个 8 连通特征及其 4 连通背景之间的差异。

给定两幅二值图像 A 和 B，组合(NOT A) AND B 会生成一幅包含出现在图 B 中而不出现在图 A 中的像素的图像。这和 NOT (A AND B 的结果不同，后者会选中在 A 和 B 中都不是“开”的像素。它也和 A AND (NOT B)的结果不同，如图 8.3 所示。算子应用的顺序很重要，而使用括号来说明运算顺序和范围则至关重要。实际上，上面列出的 4 种运算是冗余的。三种运算就足以产生所有这些结果。因此，有些系统会忽略其中一种运算（通常是 XOR）。但为清晰起见，此处所示的例子中使用了所有 4 种运算。

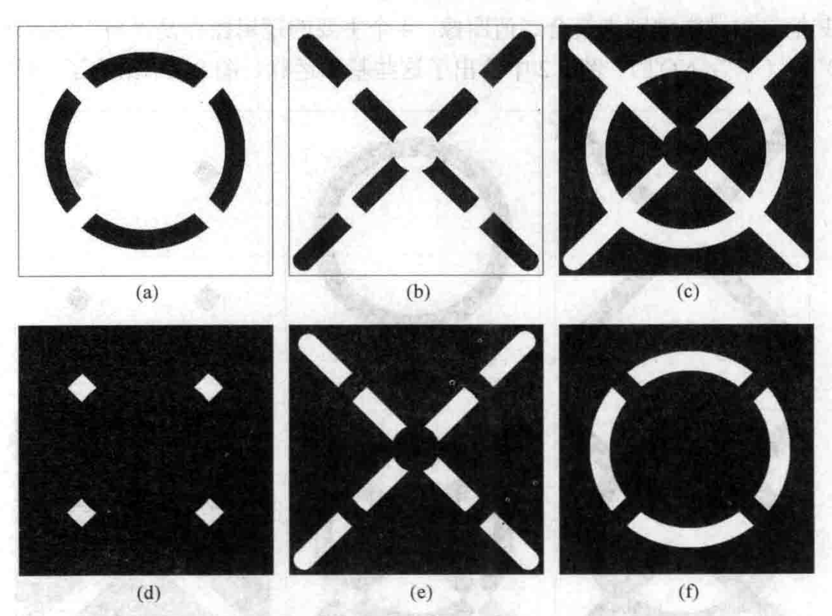


图 8.3　组合的布尔运算：(a) (NOT A) AND B；(b) A AND (NOT B)；(c) (NOT A) AND (NOT B)；(d) NOT (A AND B)；(e) (NOT A) OR B；(f) A OR (NOT B)

8.2 组合布尔运算

当用于选取前景像素的规则有多个时，可使用布尔组合来合并这些规则。最常见的情形是多波段或多通道图像，例如由卫星生成的图像。图 8.4 显示了一个例子。

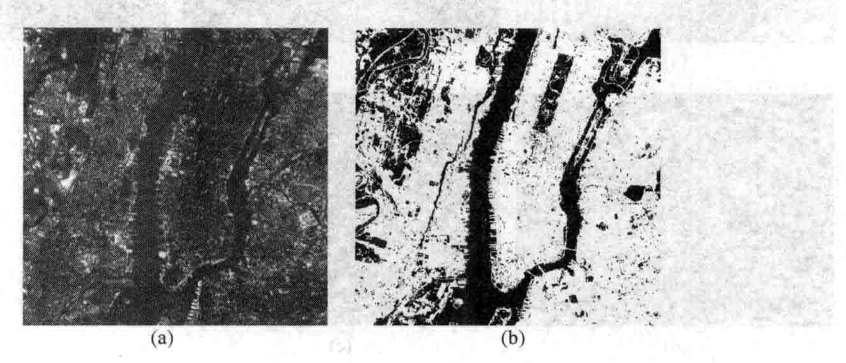


图 8.4 (a)纽约市的卫星图像；(b)（对可见光通道阈值处理得到的）水与（对红外通道阈值处理得到的）植被进行“或”运算的组合结果

另一个多通道情形出现在 SEM 中，此时常用一个 X 射线探测器来创建多幅图像（通常称为 X 射线点图），每幅图像显示一种选定元素的空间分布。这些图像中可能会因为有大量的噪声（见第 4 章）而难以进行阈值处理（见第 7 章）。但延长图像的获取时间或对图像进行处理后，可对这些图像进行阈值处理，生成可指示某些选取阈值之上元素集中位置的二值图像。

由于样本中电子速度的降低会导致有限级别的背景信号，因此通常通过将测量的 X 射线强度与高于背景几个标准差的一个阈值进行比较，来进行这一选择。该现象的物理背景在这里并不重要。点图的较差统计特性会使得我们很难直接指定一个集中度电平作为阈值。图像中一部分的 X 射线强度随不同区域变化的原因如下：①该元素集中度的改变；②选择性吸收或散射第一种元素辐射的另一种元素集中度的改变；③标本密度或表面方向的改变。由于无法保证精确地再现设备条件，因此样本间的比较存在困难。这些影响都会使得元素集中度和所记录亮度值之间的关系变得复杂。

此外，（用电子束生成 X 射线的低效性所致的）非常差的图像统计特性，意味着通常需要对这些图像进行处理，即要么通过平滑操作将图像变为灰度图像，要么使用下面介绍的形态学工具将图像变为二值图像。下面的讨论假定图像是二值图像，每幅图像都显示某种元素有意义集中度的空间分布。

如图 8.5 所示，SEM 也会生成更传统的图像。反向散射电子图像对结构给出了更好的定义，但失去了元素的一些专有属性。这些二值图像可以与 X 射线或元素信息合并，如图 8.6 所示。通过平滑点图，并对灰度图像进行阈值处理，可得到铁元素（Fe）和硅元素（Si）的 X 射线图。注意，在灰度图像中［见图 8.5(c)］，两个不同区域中铁元素的 X 射线亮度级刚好存在可区分的差异。要进行可靠的阈值处理，这一差异仍太小。甚至硅元素亮度［见图 8.5(b)］中的较大差异也难以分离。然而，布尔逻辑很容易组合图像来生成含铁但不含硅的一幅区域图像。

在反向散射电子图像中，含银区域通常更亮，但某些其他区域也较亮。另一方面，由于较差的统计特性，银元素 X 射线图没有精确的区域边界。使用 AND 运算来合并两幅二值图像，生成了想要的区域［见图 8.6(d)］。也可使用顺序更复杂的布尔逻辑运算来得到想要的结果［图 8.6(e)显示了一个例子］。

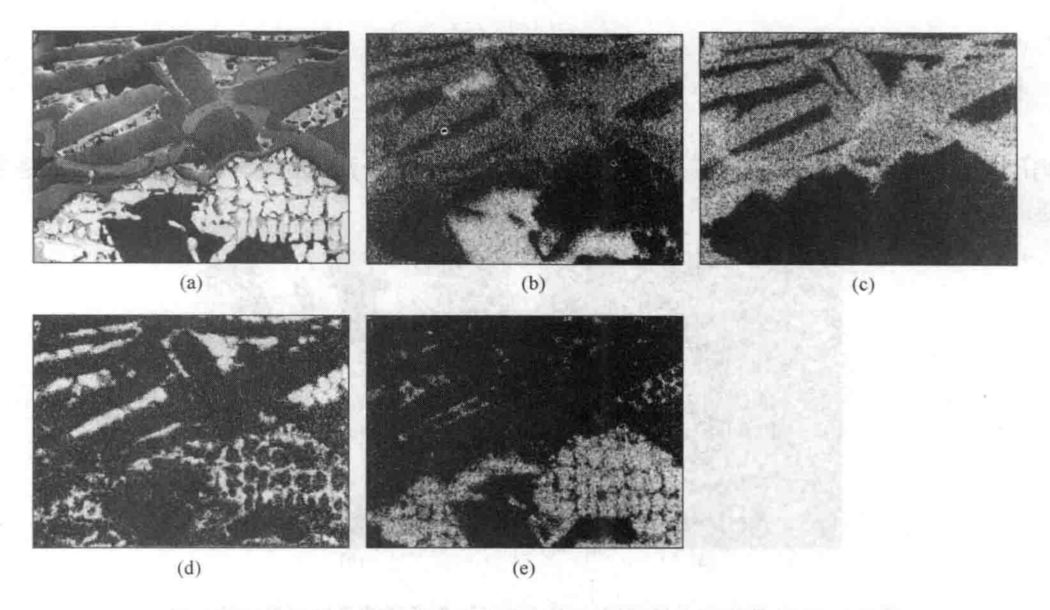


图 8.5 矿物的 SEM 结果：(a)反向散射电子；(b)硅元素 X 射线图；(c)铁元素 X 射线图；(d)铜元素 X 射线图；(e)银元素 X 射线

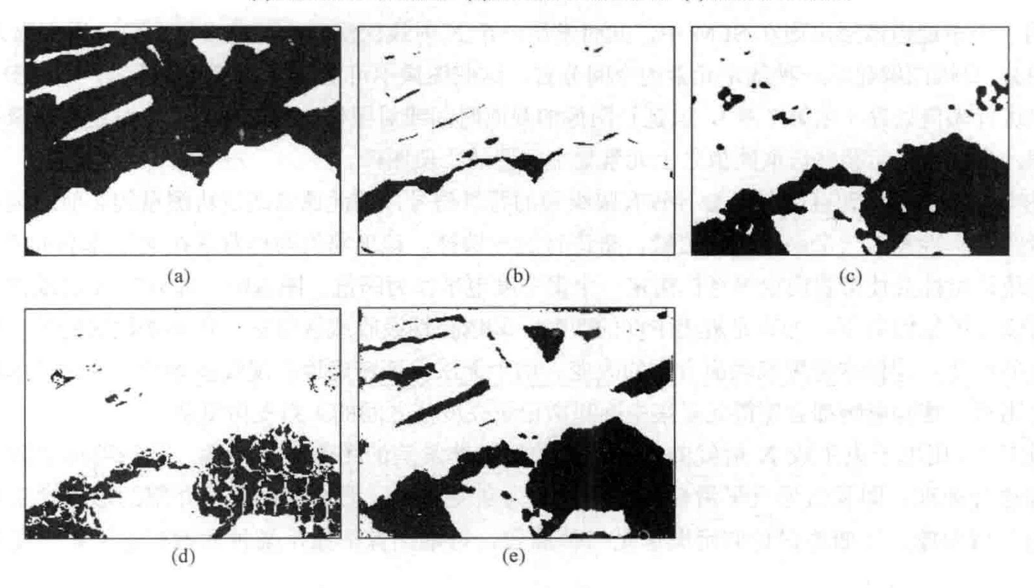


图 8.6 图 8.5 所示图像经阈值处理后的二值图像及其布尔组合：(a)Fe; (b)Fe AND NOT Si; (c)Ag; (d)Ag AND 反向散射电子图像的亮度级; (e) (Cu OR Ag) AND NOT Fe

处理含有许多元素的复杂标本也很简单。由各种成分组成的颜料粒子就是这样的一个例子。为了统计或测量某类粒子（如与增白剂和填充剂相对的染料），有必要指定那些含铁、铬或铝而不含钛或硫的粒子。这可写为

(Fe OR Cr OR Al) AND (NOT (Ti OR S))

生成的图像然后可与（二次或反向散射电子图像经阈值处理后生成的）一幅更高分辨率的二值图像组合，进而描绘出粒子的边界。多个信号的类似组合也出现在基于多重染色或荧光分子的生物样本中，以及遥感影像和地理信息系统中。执行多个运算时，要求保持组合顺序的简单性，但这并不困难。

前几章中给出的使用多个通道（如不同颜色或元素）或不同处理运算（如合并亮度值和纹理）的多数

例子中，使用了一个布尔“与”运算来组合各个单独经阈值处理后的二值图像。如上所述，“与”运算要求像素满足所有条件后才能保留。有些情形下布尔“或”运算更为合适。第 5 章中的图 5.60 就是这样一个例子。这是一幅通过偏振镜观察到的砂岩中的沙粒图像。分析仪的每次旋转都会导致不同的晶粒变亮或显出彩色。第 5 章表明，随着分析仪的旋转，在每个位置保留最亮的像素值生成了一幅显示有所有晶粒的图像。

图 8.7 给出了对同一问题的另一种解法。对每幅图像单独进行阈值处理，选取那些针对分析仪的某个旋转角度的明亮晶粒。然后使用布尔“或”运算来合并所有二值图像。所得组合结果中显示了大多数晶粒，但它要比相同数量的分析仪旋转的灰度级运算结果差一些。通常，只要可能，都要在阈值处理前执行包括多幅图像组合的图像处理，因为与阈值处理后保留的像素亮度值相比，原始的像素亮度值可提供更多的信息。

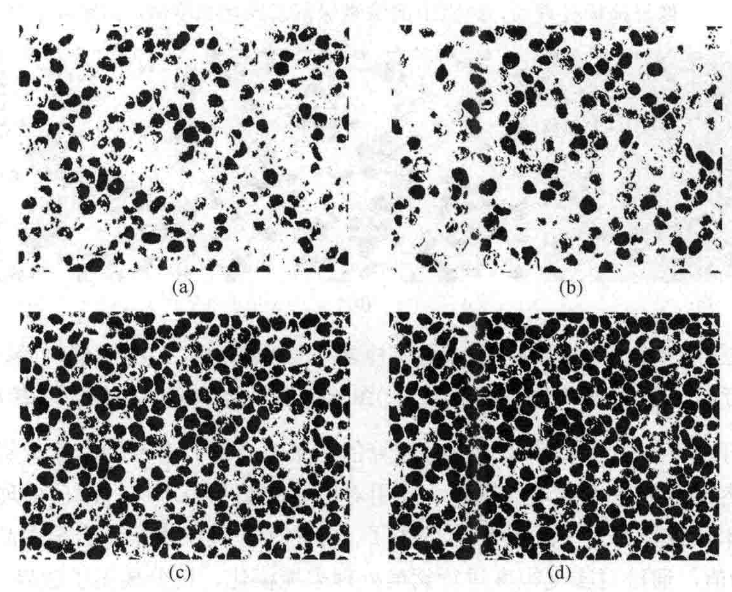


图 8.7　组合多幅二值图像：(a)和(b)对砂岩岩相切片的两幅极化光图像（第 5 章中的图 5.60）进行阈值处理后得到的图像；(c)分析仪旋转至不同角度获得的 6 幅图像的“或”运算结果；(d)阈值处理得到的二值图像，与相同的 6 幅彩色图像经组合操作得到的灰度图像（在每个位置保留最亮像素）的比较

8.3　掩模

上面使用布尔逻辑来组合图像的说明，假定了两幅图像都是二值图像（即黑白图像）。也可以使用二值图像作为掩模来修改一幅灰度或彩色图像。这通常是通过删除连续色调图像中的某部分（即设为背景）来实现的，例如创建仅显示某些感兴趣区域的图像，或选择将测量其亮度、颜色、密度等的区域。图 8.8 给出了一个例子（蛋白质分离凝胶），其中对暗点进行了阈值处理，然后将所得二值图像用做掩模来产生要测量的各个保留其原始亮度值的分离特征。图 8.9 显示了应用到彩色图像（显示细胞壁及淀粉颗粒的染色马铃薯切片）的一个类似操作。对绿色通道进行阈值处理，生成了用于颗粒的一个二值掩模，然后对彩色图像应用该掩模来隔离出颗粒进行分析。

执行这种运算的物理方法有几种。二值掩模可用在叠加或阿尔法通道中来选取要呈现的像素，或用于修改存储的图像。这是通过将灰度图像乘以二值图像，并约定二值图像在每个像素点的值为 0（“闭”）或 1（“开”）来实现的。组合灰度图像和二值图像来保留较暗或较的值，也可得到相同的结果。例如，如果对于背景像素掩模是白色的（值为 255），而对于前景像素是黑色的（值为 0），那么在每个更亮像素值的位置上删除所有背景像素，并为前景像素保留灰度或彩色值。

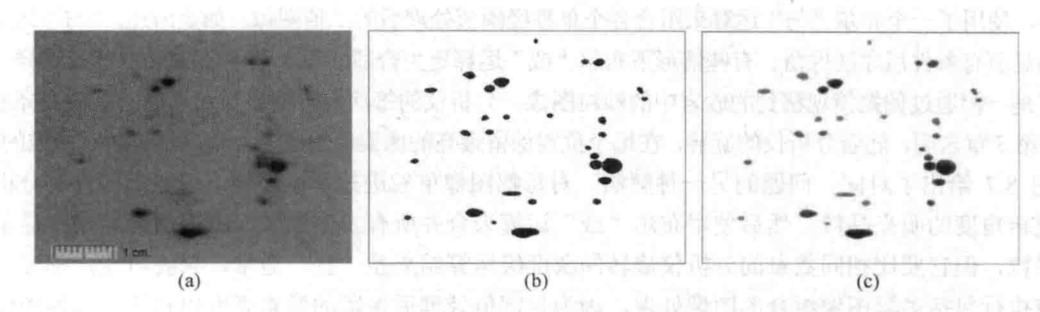


图 8.8　保留特征亮度值：(a)原始的二维凝胶；(b)经阈值处理后的斑点；(c)图像经掩模处理后，斑点中的像素保留了原始亮度值，但删除了背景

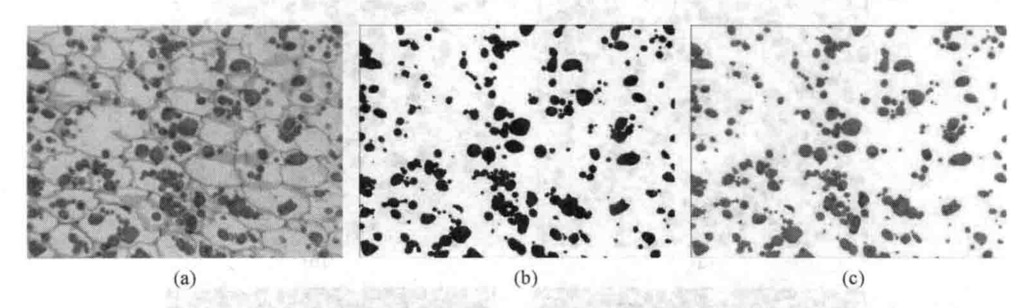


图 8.9　应用到彩色图像的掩模：(a)染色马铃薯切片的原始图像；(b)对图(a)的绿色通道进行阈值处理后得到的二值图像；(c)图(b)中的二值掩模应用到原始图像后的结果

这种能力常用于显示不同处理和阈值处理运算的结果。查看具有原始灰度或彩色信息的所选像素而非二值图像，更容易判断阈值处理的性能。使用对图像的某个版本进行阈值处理得到的掩模，来查看图像的另一个版本，也是有用的。图 8.10 给出了一个例子，其中表示铝合金中晶粒边界方向角（由 Sobel 求导得到）的值，通过对梯度幅度进行阈值处理来掩模化，仅分离出了边界。

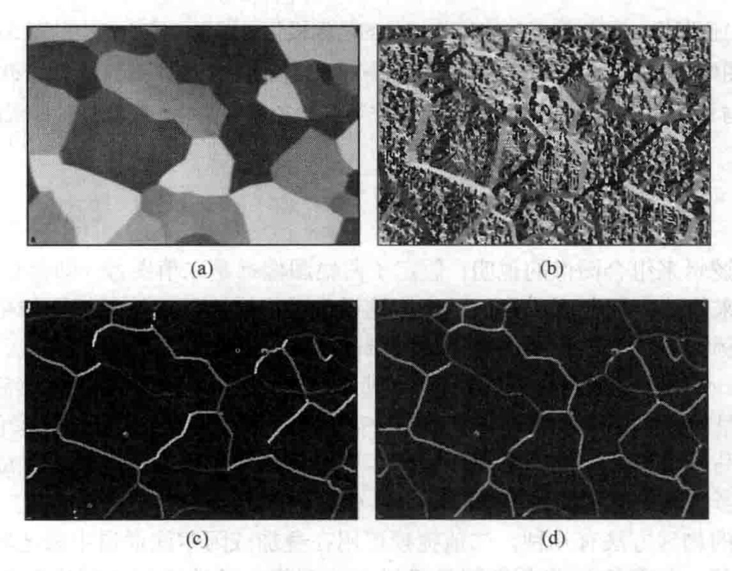


图 8.10　用一幅图像对另一幅图像进行掩模处理。仅在具有较大梯度幅度的区域中，显示了应用到铝合金光学显微图像的 Sobel 梯度方向：(a)原始图像；(b)图像中所有点的方向值（灰度）；(c)仅显示沿具有较大梯度幅度边界的值的掩模处理结果；(d)用于沿颗粒边界向方向值赋色的彩色 LUT

图 8.11 显示了另一个应用掩模处理和布尔图像组合的例子。尽管本质上这是一种修饰，但它仍然是有用的，且被广泛应用。如果图像中包含了全范围的亮度值，那么叠加在图像上的黑色或白色标签可能难以阅读。在该例中，通过膨胀操作（在下面介绍），标签用于创建一个在所有方向上均大 1 个像素的掩模。然后在写入黑色标签之前，使用该掩模将灰度图像中的像素擦成白色，反之亦然。结果保证了标签的易读性，但同时图像出现了少量模糊。

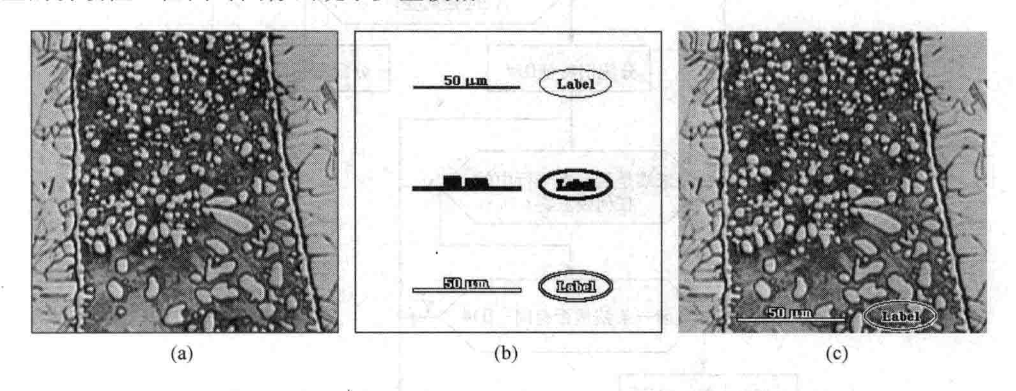


图 8.11　使用掩模将标签添加到一幅图像中：(a)含有白色和黑色区域的原始图像，看不到简单叠加的文字。(b)对标签进行膨胀操作，并与原始图像进行 XOR 运算所生成的掩模；(c)所得结果与灰度图像叠加后的效果

8.4　从像素到特征

上面描述的布尔运算处理的是图像中的各个像素。出于很多目的，特别是为了测量，需要识别出形成连通整体一部分的像素。如第 7 章所述，对单个特征中的像素，可以对 8 连通或 4 连通相接采用一种约定（有时称为大对象，以表明相连像素组未被推断为图像中的任何特定部分）。不管采用哪种约定，将像素分组到各特征中都是重要的一步（Levialdi，1992；Ritter，1996）。

识别所有相连像素组的一种方法是从一个像素（随机选取的任意一个“开”像素）开始，并检查其 4 邻域或 8 邻域位置，把每个“开”像素标为同一特征中的一部分，接着重复该运算直到没有相邻像素保留。然后选择一个新的未标记像素，并重复该运算，直到图像中的所有“开”像素都被标为某个特征的一部分。处理这种深度递归的“种子填充”运算的常用方法是，创建一个栈，当像素是已标记像素的相邻像素时，则将其位置保存到栈中。检查这些像素的相邻像素时，则将它们从栈中删除。当栈为空且所有相接的像素已被定位并识别时，处理终止。

应用到灰度或彩色图像的相同方法，会检查每个相邻像素查测试这些像素值是否在某个阈值范围内。这种区域生长的结果是一个选择性的图像分割，如第 7 章中的图 7.52 所示。

按组处理像素效率更高。如果图像如第 7 章中讨论的那样已经进行了游程编码或弦编码，那么可知弦内的所有像素是相接的，与它们中的任意一个相接等价于与全部像素相接，而用于相接的候选像素是相邻线上方或下方的那些像素。这一事实就使得开发出一种一次遍历图像的简单标记算法成为可能。将每条弦的端点与前一行中的弦的端点进行比较；如果它们相接，则将前一行的标签贴到这条弦上。否则，使用一个新的标签。

如果一条弦与前一行中具有不同标签的两条弦相接，则这两个标签相互识别（例如，这会处理字母 U 的底部）。某个标签的所有出现可变为另一个标签的所有出现。当遍历图像或完成弦表时，所有弦即所有像素都会被识别，因此可知标签（亦即特征）的总数。图 8.12 用流程图的形式显示了这种算法。

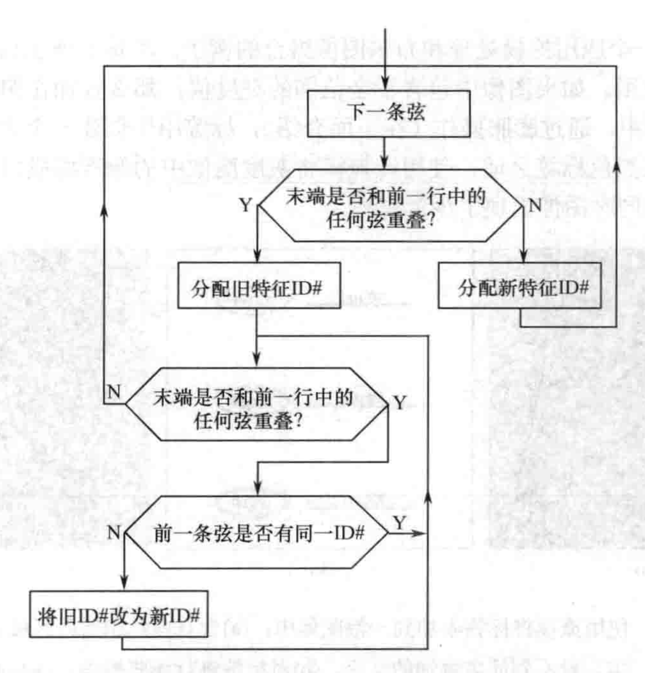


图 8.12　将游程编码或弦编码阵列中的相接像素组合到特征中并分配 ID 号的流程图

对于边界表示（包括链码），由于边界已经表示了围绕特征的一条闭合路径，因此分析是部分完整的。如果特征中不包含孔洞，且没有特征被其他包围，那么它可以提供完整的信息。遗憾的是，情形并非总是如此。通常需要重建像素阵列来以特征标签识别像素（Kim et al., 1988）。

在任何情形下，只要标记了各个特征，就可以进行几种其他的布尔运算。一种运算是找到并填充特征内的孔洞。属于孔洞一部分的任何像素都是“闭”像素（即背景的一部分），但它们被“开”像素包围着。对于边界表示，这意味着像素位于边界内。对于像素表示，这意味着像素不与最终形成一条到视野边缘的路径的其他像素相连。

回顾相接的约定（8 连通或 4 连通）可知背景和前景必须存在不同，反转图像（黑白互换）并对所得像素进行标记，就如同它们是特征一样，可很容易识别出孔洞，只不过使用的是相反的连通逻辑，如图 8.13 中的各步所示。反转图像中的特征是原始图像中的孔洞。如果将这些像素加回到原始图像中（使用布尔“或”运算），就填充了原始特征中的任何内部孔洞。

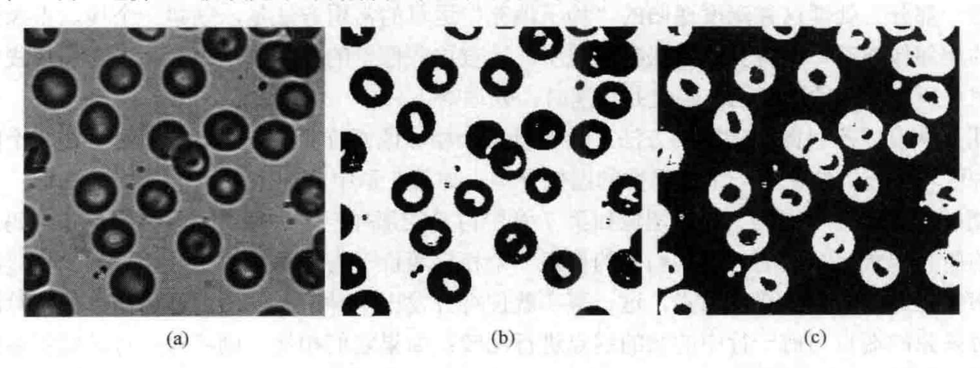


图 8.13　血液红细胞的光学显微镜图像：(a)原始图像；(b)经阈值处理后的图像，显示了血细胞的较厚外部边缘，但未显示出较小的中心区域；(c)图(b)反转对比度后的结果；(d)从图(c)中删除边缘相接背景后保留了孔洞；(e)组合图(d)中孔洞和图(b)经布尔“或”运算找出的特征后的效果；(f)删除小特征（灰尘）、(无法测量的）边缘相接特征及隔离出图(e)相接特征后的结果；(g)叠加到原始图像上后，圈出了可被测量的血液红细胞

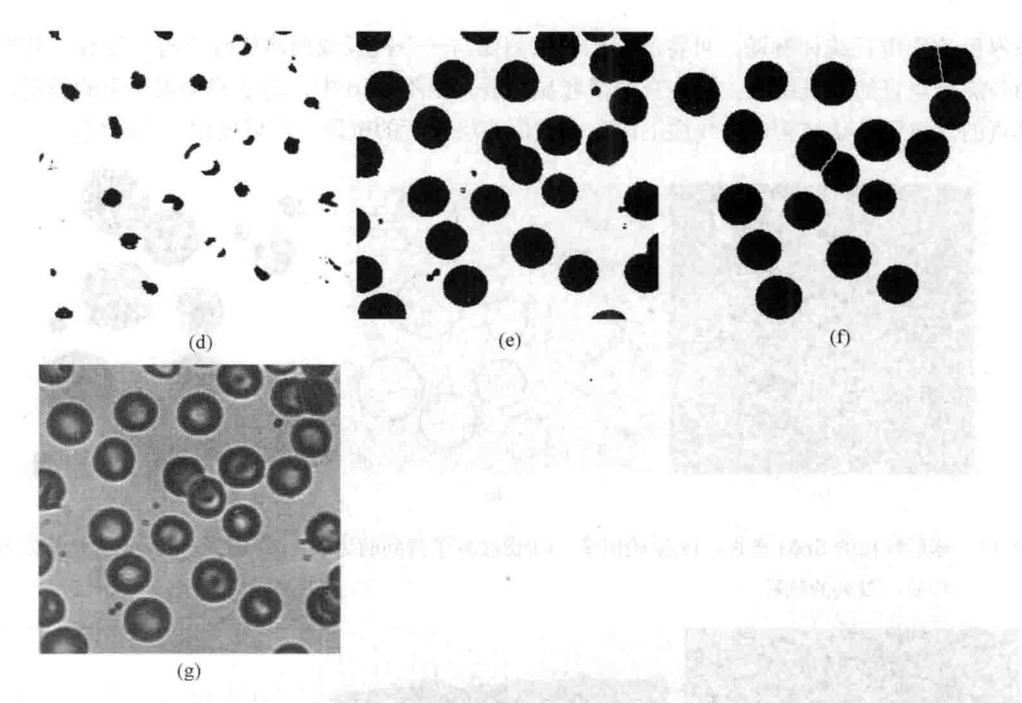


图 8.13　血液红细胞的光学显微镜图像：(a)原始图像；(b)经阈值处理后的图像，显示了血细胞的较厚外部边缘，但未显示出较小的中心区域；(c)图(b)反转对比度后的结果；(d)从图(c)中删除边缘相接背景后保留了孔洞；(e)组合图(d)中孔洞和图(b)经布尔“或”运算找出的特征后的效果；(f)删除小特征（灰尘）、（无法测量的）边缘相接特征及隔离出图(e)相接特征后的结果；(g)叠加到原始图像上后，圈出了可被测量的血液红细胞（续）

应用这种技术的另一个例子如图 8.14 所示，它是搪瓷表层中球形孔隙的一幅显微镜图像。许多小孔中心处的亮点的亮度会随孔深而变化。必须在图像经阈值处理后得到的二值图像中填充特征来修正它们。注意，由于软件无法知道边缘之外的内容，且无法知道特征是否闭合，因此并未填充与图像边缘相交的小孔中的孔洞。但无论怎样，因为同样的原因，这些小孔无法测量。

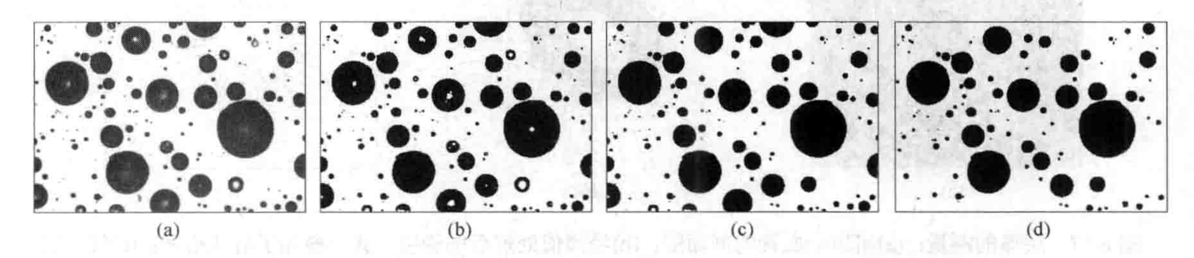


图 8.14　铁上搪瓷护膜抛光截面的光学显微镜图像：(a)反射光的亮点出现在许多小孔的内部（取决于它们的深度）；(b)阈值处理后的图像中保留了小孔内的孔洞；(c)填充特征内部的孔洞所创建的小孔图像；(d)删除边缘相接特征仅保留了可测量的小孔

空洞填充常和边缘查找一起使用。第 5 章中的算法勾画出了物体的边界，但对于许多类型的测量，内部像素也必须为“开”像素。SEM 图像通常需要这种类型的处理，如图 8.15 中的例子所示。亮度会随局部表面斜率变化，所以较亮边缘和较暗中心会频繁出现颗粒。

对图像应用测量模板时，布尔“与”运算特别有用。例如，考虑测量线缆或平板的涂层厚度。在图 8.16 和图 8.17 的例子中，对涂层可以很容易地进行阈值处理，但涂层的厚度并不均匀。为了得到一

组离散厚度值来进行统计解译，可将涂层的二值图像与一个模板或网格进行“与”运算，其中模板或网格由与涂层垂直的直线组成。这些直线很容易测量。在图 8.16 中，对于平坦表面上的涂层，这些直线是垂直的。对于柱状结构，如线缆上的一个类似涂层或管的壁厚，可以使用一组射线。

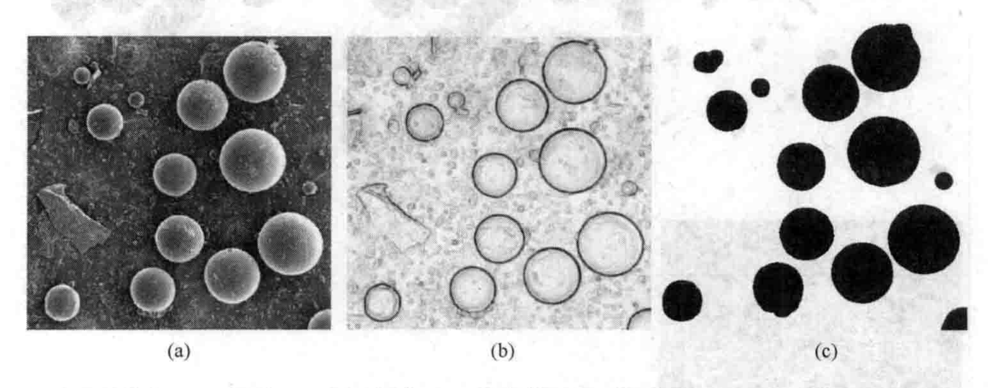


图 8.15 球形颗粒的 SEM 图像：(a)原始图像；(b)梯度算子得到的边缘；(c)对边缘进行阈值处理并填充中心后，得到的颗粒

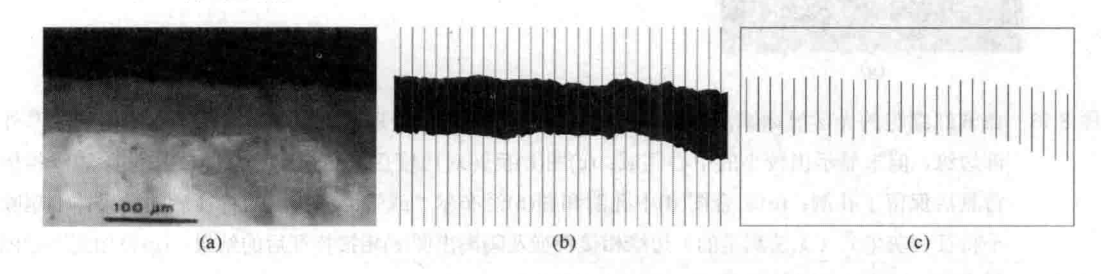


图 8.16 涂层厚度的测量：(a)漆层的剖面图；(b)经阈值处理后的涂层，其上叠加了垂直线格网；(c)直线与涂层“与”操作，产生了可测量的线段

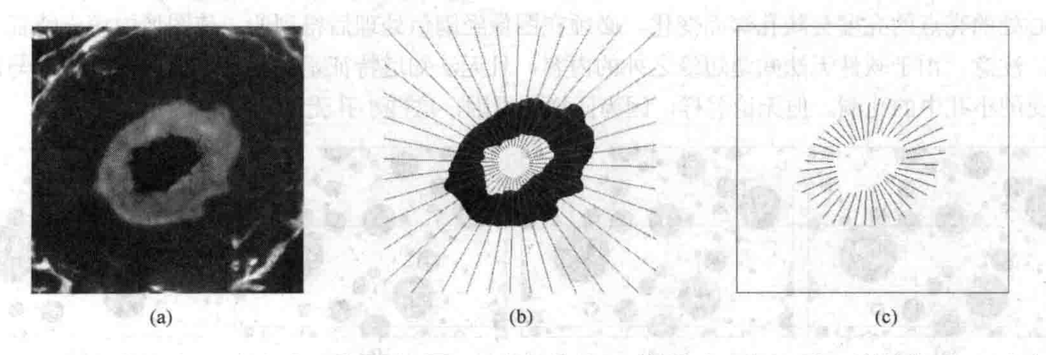


图 8.17 层厚的测量：(a)组织中血管的剖面图；(b)经阈值处理后的管壁，其上叠加了射线格网；(c)线与层“与”操作生成了可测量的线段（注意网格线与管壁不垂直对齐时引入了误差，它并不恰好是圆形）

在图 8.17 的例子中，血管在剖面中近似为圆形，而且直线与管壁并不垂直相交，因此引入了不可接受的测量误差。测量不规则或弯曲层宽度的另一种方法将在下面介绍。非圆形横剖面还表明，剖面与血管轴不垂直，因此在分析中又引入了另一种误差。对二维剖面图像进行三维结构测量，是通过体视学技术来处理的，这将在第 9 章详细讨论。

选择合适的网格对于测量成功至关重要。第 9 章中将介绍对结构进行主要体视测量的技术，以求出所现成分的体积、面积、长度和拓扑特性。许多此类过程的执行方法是，统计各网格与感兴趣结构的交点数量。网格通常由点阵或线阵组成。所用到的线包括直线、圆弧和摆线的规则及随机网

格，具体取决于期望的测量类型、用于选择并准备待成像标本的过程，以及用于生成检查用表面的剖面方向。在所有情形下，如果图像经阈值处理后可成功地勾画出结构，那么与适当的网格进行布尔“与”运算，就可生成可测量的结果。在有些情形下需要测量线长，但许多情形下只需统计由网格产生的交点数。

甚至对于非常复杂图像（即自动处理和阈值处理不能描绘出感兴趣结构的图像），把网格作为掩模进行叠加可能十分重要。许多只需要统计不同类型网格和感兴趣特征的交点数的体视过程非常高效，而且可以提供有价值的结构参数的无偏估计。通过覆盖合适的网格——点阵或线阵，后者包括直线、圆环和摆线——来组合原始图像与处理后的图像，可增强结构的可见性，使用户能识别出重要的特征和交点（Russ，1995a）。计数可以手工进行，也可通过计算鼠标单击的次数或统计用户在图像上放置的标记数量，由计算机辅助完成。结合人工识别与计算机辅助方法来获取、处理及显示图像，并生成合适的网格，为许多图像分析问题提供了有效的解决方案。

8.5 特征的布尔逻辑运算

识别并标记像素进而组合为特征后，可在特征而非像素层面执行布尔逻辑操作。图 8.18 显示了根据特征进行“与”运算的原理。如果特征的任何部分存在相接，则保留整个特征，而不只是保留两幅图像中共有的像素。这保留了整个特征，以便在第二幅图像中选取它时可被正确地计数或测量。

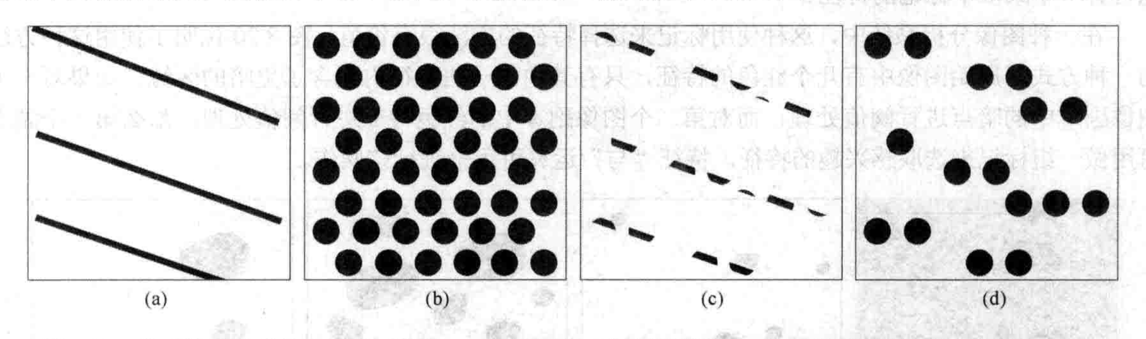


图 8.18 基于特征的“与”运算示意图：(a)标记图像；(b)测试图像；(c)图(a)和图(b)的基于像素的布尔“与”运算结果；(d)基于特征的“与”运算结果，其中图(a)中的标记选取了图(b)中的标记

这种方法又称为“基于标记”方法，它使用一幅图像中的像素或像素组作为标记，来选取第二幅图像中的特征。有效的实现方法是，至少对待组合的一幅图像应用特征标记运算。如前所述，一幅图像中的相接像素被识别为特征。因此，将一幅图像内某个特征中的每个“开”像素与第二幅图像中的对应像素进行核对。如果该特征中的任意像素与第二幅图像中的“开”像素匹配，那么第一幅图像中的这个特征就会复制到结果中。

这不是唯一可能的实现。同样可以将第二幅图像中的每个像素与第一幅图像中的对应像素进行核对，但效率会低一些。给出的方法将比较限制在“开”像素范围内，且对于任何特征，只要其中的任意像素匹配，检验就会停止。可实现同样结果的另一种方法是，对标记进行膨胀（在本章后面介绍），但只允许膨胀在某个（由第二幅图像中的特征定义的）掩模内进行。这种方法比较慢（因为它是迭代进行的），且对于具有非常复杂或扭曲形状的特征，要到达不再改变的终点，可能需要很多次迭代。

与更常见的基于像素的“与”运算不同，标记选择并不满足交换率。这表明（A 特征 AND B）与（B 特征 AND A）的结果并不相同，如图 8.19 所示。特征“与”运算和“非”运算一起使用很容易实

现，例如执行同样的过程并删除第一幅图像中与第二幅图像中任意像素匹配的特征。不需要对特征进行“或”运算，因为这种“或”运算的结果，与基于像素的常规“或”运算结果相同。

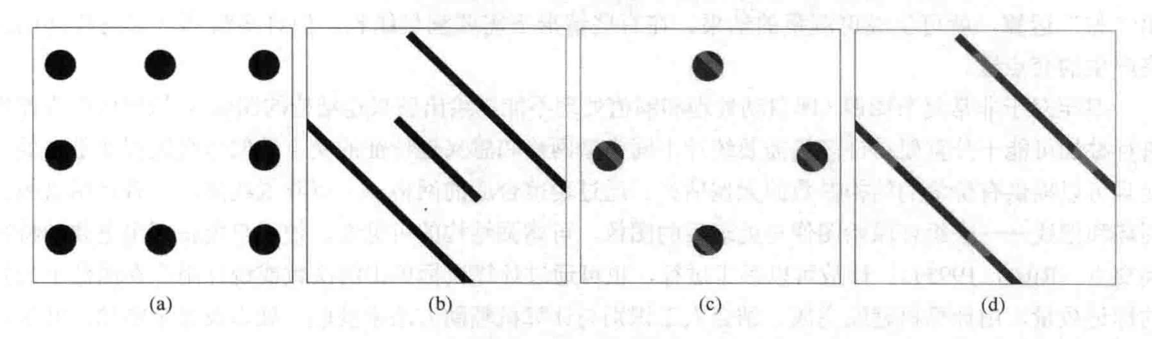


图 8.19　(a)和(b)用于进行基于特征的布尔逻辑运算的两幅测试图像；(c)使用图(b)作为标记来选取图(a)中的特征；(d)使用图(a)作为标记来选择图(b)中的特征。在图(c)和(d)中，像素“与”运算保留的公共像素显示为彩色

特征“与”运算的一种主要用途是，使用特征中的标记来选择它们。例如，它们可能是包含有染色细胞器的细胞，或是包含有一个特征核的合成纤维。在任何情形下，两幅二值图像由阈值处理产生。在一幅图像中勾画出了所有特征，而在另一幅图像中则定义了标记。应用特征“与”逻辑，然后选取包含有一个或多个标记的特征。

在一种图像分析系统中，这种使用标记来选择特征的方法很有价值。图 8.20 说明了使用这种方法的一种方式。原始图像中有几个红色的特征，只有其中的一些特征内包含了更暗的区域。如果对一个图像副本中的暗点进行阈值处理，而对第二个图像副本中的红色特征进行阈值处理，那么第一个副本可用做一组标记来选取感兴趣的特征。特征“与”运算可用来进行该操作。

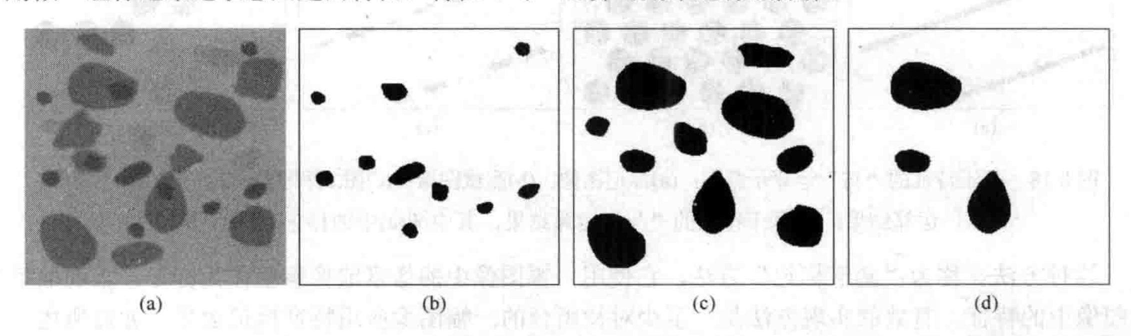


图 8.20　使用标记进行特征选择示例：(a)原始图像；(b)经阈值处理后的暗点用做标记；(c)深红色亮度经阈值处理后的特征；(d)包含暗点的那些红色特征的二值图像，它是由基于特征的“与”运算生成的

在实际应用中，选择感兴趣特征的标记图像可通过处理获得，或通过使用多通道图像中的另一个通道来获得。图 8.21 显示了一个例子。只有包含绿色细胞核的细胞被选中，为便于测量，它们是整体选中的。使用特征“与”性能的一个相关过程是成核剂的使用（Gundersen et al., 1988），成核剂是按照细胞中出现的唯一标记（如细胞核）来选择并统计组织切片中细胞的一种体视学工具。

图 8.22 显示了使用标记来选择物体的另一个例子。糖果图像经阈值处理后，内部出现了由字母 m 导致的一些孔洞。反转图像并删除与边缘相接背景后，保留这些孔洞作为标记。对原始二值图像填充孔洞并应用分水岭分割（在下面描述）后，可用这些标记来仅选择最初包含有孔洞的那些特征。使用结果图像作为掩模，并应用到原始图像后，仅显示了其上带有字母的糖果。

在非常不同的尺度上，这种方法可用在航拍图像中，以选择、统计或测量包含任意建筑物的建

筑用地或包含动物的旷野。该技术也可用在 X 射线图像中，以选择 SEM 图像中的粒子，前提是这些 X 射线信号源自对 X 射线检测器可见的这部分粒子。只要其任何部分产生了可识别的 X 射线信息，就保留整幅粒子图像。

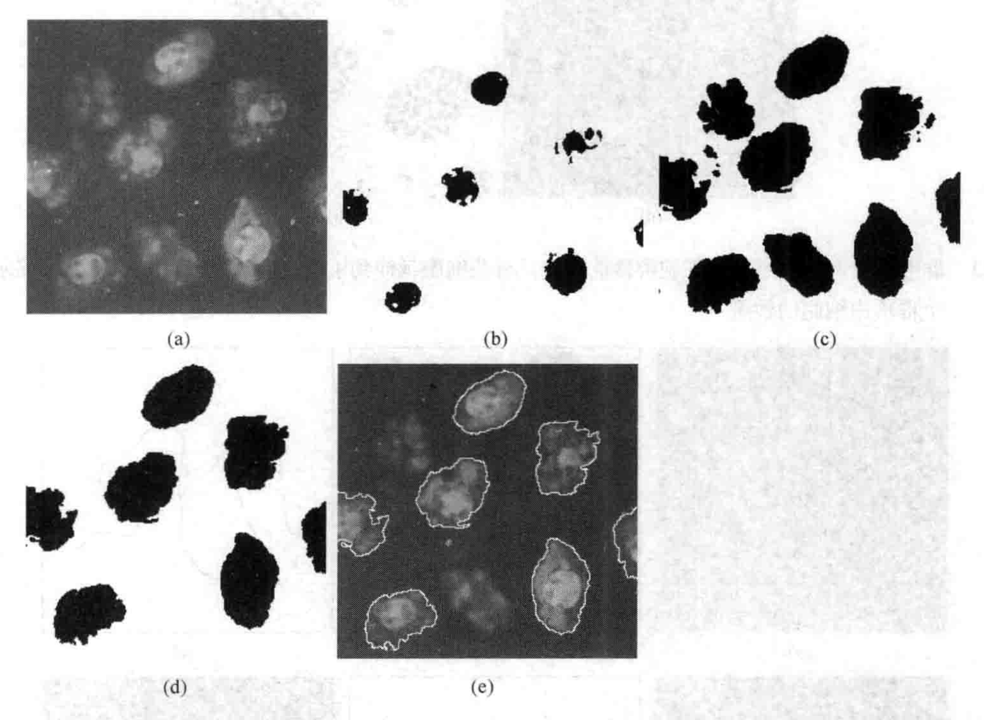


图 8.21 标记选择的应用：(a)具有染色细胞核的细胞的原始图像；(b)基于绿色亮度进行阈值处理后的细胞核；(c)基于红色亮度进行阈值处理后的细胞；(d)特征“与”运算的结果，仅显示了含有绿色细胞核的那些细胞；(e)图(d)叠加到原始图像上后，圈出了特征

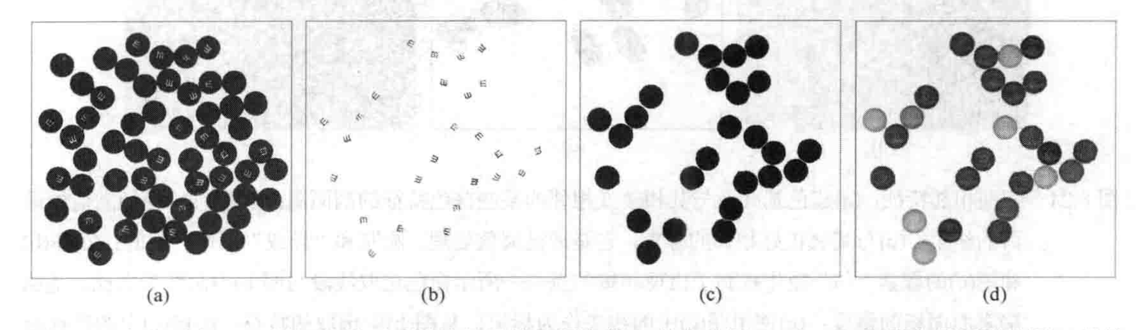


图 8.22 标记选择：(a)经阈值处理后的糖果图像（源自第 7 章中的图 7.11）；(b)印刷字母识别为图(a)的特征中的孔洞；(c)经孔洞填充和分水岭分割后，使用图(b)作为标记来选取图(a)中的特征；(d)将图(c)作为掩模应用到原始图像，结果中仅显示了字母 m 朝上的糖果

标记选择也可用于隔离位于或部分位于某些区域内的特征。例如，在图 8.23 中，菌落包含有将被计数和测量的细菌细胞，尽管有些细胞延伸到了菌落边界之外。特征“与”逻辑运算可将它们分配到合适的菌落并计数，但在它们离开或重新进入该区域时则不重复计数。

图 8.24 显示了一种常见的情形，其中将要选取的只是与一条边界相接的物体。在图像序列中，生成了区域的轮廓并被用做一个标记来选取相接的特征。用像素“与”运算将边界线和特征组合，可找到相接出现的位置，且这些线段的总长与边界线的总长之比，就是与物体相接表面面积部分的一个测度。使用边界线作为标记，可选取相接物体，进而对其计数或测量。

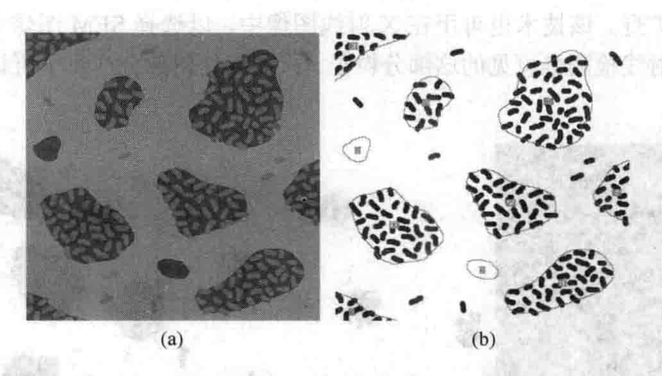


图 8.23　群落计数：(a)表示细菌细胞群落的图像，有些细胞延伸到了着色区域之外；(b)计算结果显示了每个群落中细胞的数量

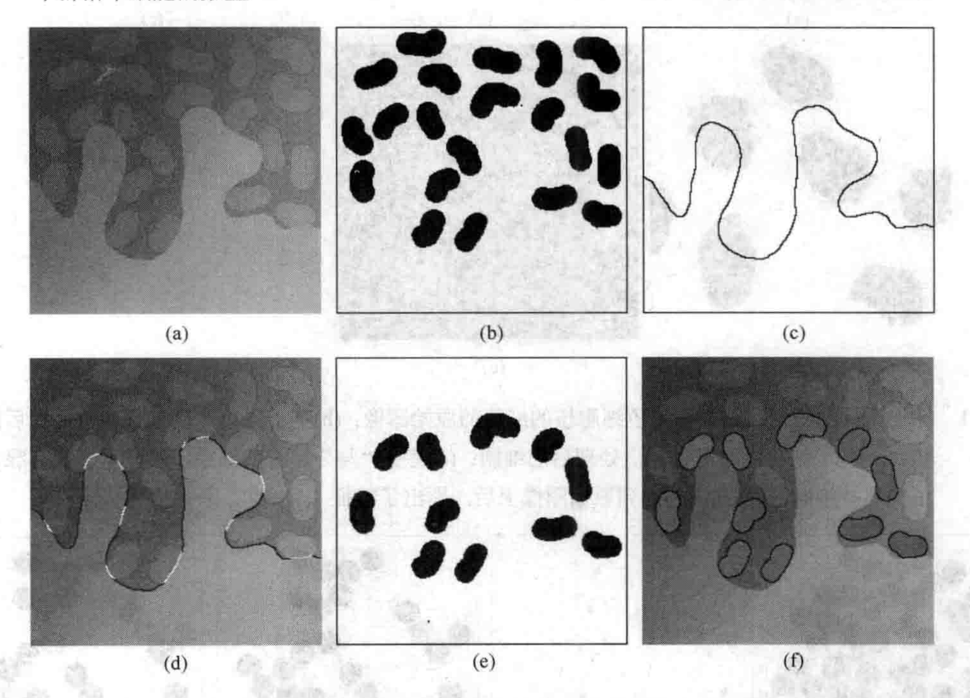


图 8.24　识别相邻特征：(a)蓝色基底及与其相接或相邻的某些橙色特征的剖面图；(b)橙色特征经阈值处理后的结果；(c)与基底正好相邻的像素，它是通过阈值处理、膨胀和“异或”操作产生的；(d)图(b)和图(c)的像素“与”操作找到了出现相接的点——所示白色选取线段与图(b)中总线长之比，是相接表面面积的测度；(e)使用图(c)中的线条作为标记，从图(b)中选取的特征；(f)图(e)中的特征叠加到原始图像上后识别的特征

8.6　按位置选取特征

图 8.24 中所示边界交叉或边界相接特征的识别方法的一种推广是，使用标记选择来识别离线条或区域在指定距离内的特征。下面介绍的膨胀，会将线条转换为指定宽度的条带，或将区域在各方向上均匀地扩大。线条，如颗粒的边界或显微镜下看到的细胞壁，或卫星图像中的一条公路，经处理后均可生成具有所选宽度的线条。使用膨胀后的线条或区域来选择相接的特征时，会选择离原始线条或边界某个距离范围内的那些特征，而不管这些特征的大小和形状。使用不同的膨胀距离重复这个运算，

可根据到不规则边界的距离来分类特征，或对特征计数。

图 8.25 显示了一个例子。红色线条和蓝色特征可分别视为颗粒边界和微粒，或高速公路网和商业中心，因为该方法可在任意尺度应用。对线条进行膨胀可提供标记，而特征“与”运算会选择在设定距离内存在任何像素的所有目标。

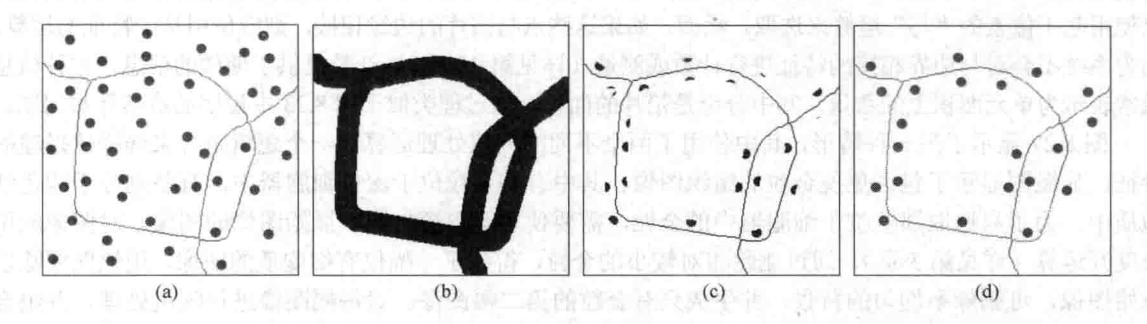


图 8.25 像素“与”运算和特征“与”运算的比较：(a)含有特征和线条的图形；(b)通过膨胀操作展宽线条；(c)图(b)与图(a)中的蓝色特征进行像素“与”运算，以选取不完整的特征和分成两部分的特征；(d)使用图(b)作为标记对图(a)中的蓝色特征进行“与”运算，选取了指定边界距离范围内的所有特征

无论是基于像素的传统布尔逻辑，还是基于特征的标记选择，都可基于位置来选取特征。图 8.26 显示了适合使用基于像素的“与”运算的一种情形。图像显示了涡轮叶片上等离子喷涂涂层的冶金截面。此类涂层上总会出现一定数量的氧化物，通常这不会有问题。但若灰色氧化物出现在了涂层－衬底的相交处，就可能使交接处断裂，或导致涂层的剥落。对图像进行阈值处理以选取氧化物，然后将结果和表示相交处的线条（这种线条的获得方式是，先对金属衬底进行阈值处理，然后使其膨胀 1 个像素，再后进行“异或”运算）进行“与”运算，直接测量出相交处的污染部分。

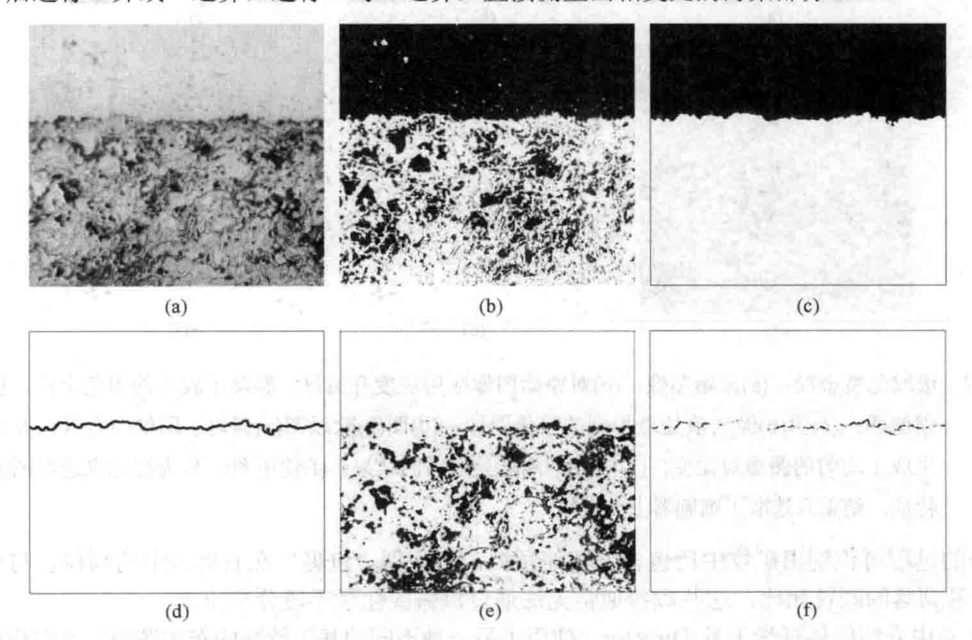


图 8.26 隔离出涂层/衬底边界中的氧化物：(a)钢上等离子喷涂涂层的显微截面灰度图像（涂层位于底部）；(b)涂层和衬底中的金属经阈值处理后的结果；(c)对图(b)应用腐蚀和膨胀后，填充了孔洞并删除了小特征，进而给出了钢衬底的表示；(d)对图(c)膨胀并与原始图像“异或”运算后得到的边界线；(e)涂层中的氧化物经阈值处理后的结果；(f)图(d)与图(b)经像素“与”运算后的结果，其中仅显示了被氧化物所占据的相交部分

使用孔径或掩模图像可将第二幅图像的分析限制到孔径内的那些区域。考虑统计箔片上的点数：这些点要么用于评估飞机喷雾操作的覆盖均匀性，要么用于评估菌点或霉点引发疾病的可能性。获取的图像是矩形的，但箔片不是矩形的。箔片的外部还有与点的亮度相似的区域。先生成箔片的二值图像，然后使用特征“与”运算将它与整个图像组合，来选择那些本身在箔片上的点。如果这些点很小，那么可以使用基于像素的“与”运算来选取。然而，如果这些点与箔片的边缘相接，则应使用基于特征的运算，因为系统不会对与边界相接的特征进行计数或测量（详见第 10 章）。计数提供了期望的信息，这种信息通常表示为单元面积上的数量，其中分母是箔片的面积。该过程类似于图 8.23 中说明的群落计数问题。

图 8.27 显示了另一种情形，其中使用了两个不同的阈值处理运算和一个逻辑组合来选择感兴趣的特征。显微图显示了包含免疫金粒的组织图像。其中有些金粒位于深色细胞器中，有些则位于浅色细胞质中。为了只选取那些位于细胞器中的金粒，需要使用两幅源自同一原始图像的图像。对图像应用灰度开运算（详见第 5 章），通过删除相对较小的金粒，得到了一幅仅有细胞质的图像。用该图像除以原始图像，可删除不均匀的背景，并生成只有金粒的第二幅图像。对每幅图像进行阈值处理，并组合经阈值处理后的两幅图像，就隔离出了最初位于细胞质上的金粒。由于基于像素的“与”运算仅保留位于细胞质边缘上的那些金粒，因此对作为标记的细胞质使用了基于特征的“与”运算来选择金粒。

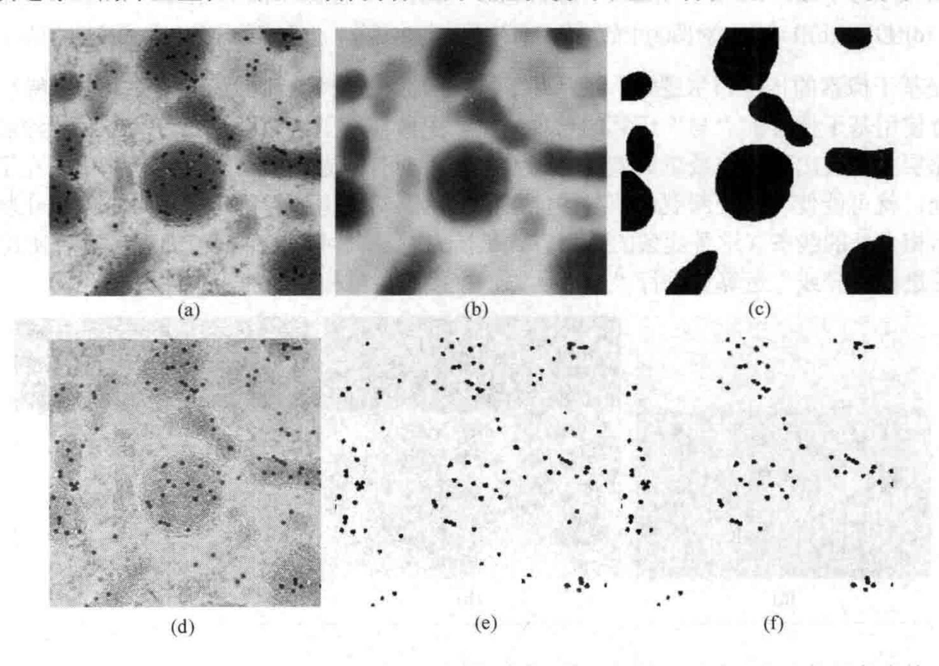


图 8.27　选取免疫金粒：(a)原始图像；(b)对原始图像应用灰度开运算，删除了较小的黑色金粒，保留了细胞器；(c)图(b)经阈值处理得到的二值图像；(d)图(b)除以图(a)调匀了图像的背景，并为金粒生成了均匀的图像对比度；(e)图(d)经阈值处理后的结果；(f)使用图(c)作为标记来选取图(e)中的特征，结果只选取了细胞器上的金粒

类似的过程可识别出矿物中所包含的其他颗粒，如识别“包裹”在石英岩中的颗粒，与那些未包裹或很容易剥落的颗粒相比，这些颗粒通常无法通过机械或化学手段分离出来。

第 9 章中介绍的体视学工具 Disector，使用了另一种不同的基于特征的布尔逻辑，来得到单位体积内特征数量的无偏和直接测量（Sterio，1984）。这种工具需要匹配两幅图像中的特征，这两幅图像表示由已知距离 T 分隔的平行平面。这些特征表示了三维物体与这些平面的相交。忽略与两个平面都相交的物体，只对仅与一个平面相交的物体计数。单位体积内物体的总数可表示为

$$N_{\mathrm{V}} = \frac{\text{数量}}{2 \cdot \text{Area} \cdot T} \tag{8.1}$$

式中，Area 表示每幅图像的面积。这种方法的优点是，它对物体的形状和大小并不敏感，但需要两个平面间的距离足够近，以便平面间不会丧失信息。实际上，这意味着距离 T 相对于任何重要的物体尺寸必须足够小。

当 T 很小时，多数物体会与两个平面相交。这些平面上的特征并不完全对应，但我们期望它们至少是部分重叠的。截取三维物体时，我们期望一个平面上的相交与第二个平面上的相交重叠，因为在 T 很小时，多数物体都会穿过两个平面，只有少数不穿过两个平面的物体才被计数，因此需要检查一个较大的图像区域来获得统计上有用的计数数量。这就使得使用基于特征逻辑的自动方法更具吸引力。

既然在两幅图像中重叠的特征不被计数，那么期望的结果就是基于特征的“异或”运算。一种使用标记选取的过程，可对那些在两个平面内未被特征“与”选取的特征计数。由于该逻辑运算不满足交换律，因此需要执行两个运算：（#1 NOT 特征-AND #2）和（#2 NOT 特征-AND #1），然后统计剩下的特征。详见图 8.28 中的示意性说明。

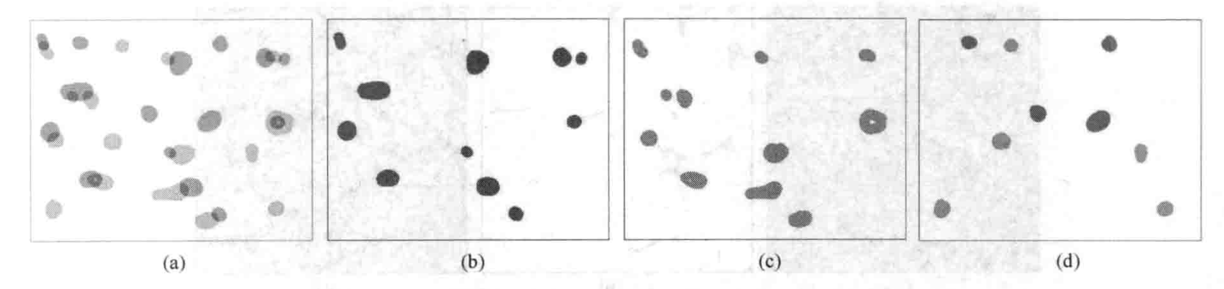


图 8.28　Disector 工具实现的图形：(a)两幅截面图像，以不同颜色显示了匹配特征的叠加；(b)（#1 特征-AND #2）选取了平面 2 中的那些与平面 1 中匹配的特征；(c)（#2 特征-AND #1）选取了在另一个平面中匹配的特征；(d)组合两个结果并删除匹配特征而只保留两个平面中不匹配特征，并对不匹配特征计数

图 8.29 显示了一个典型的应用。两幅图像是某食品上油滴的共焦光学显微成像切片，切片间的距离已知。每幅图像经阈值处理后，产生了颗粒相交的二值图像，相接特征是用分水岭方法分开的（见

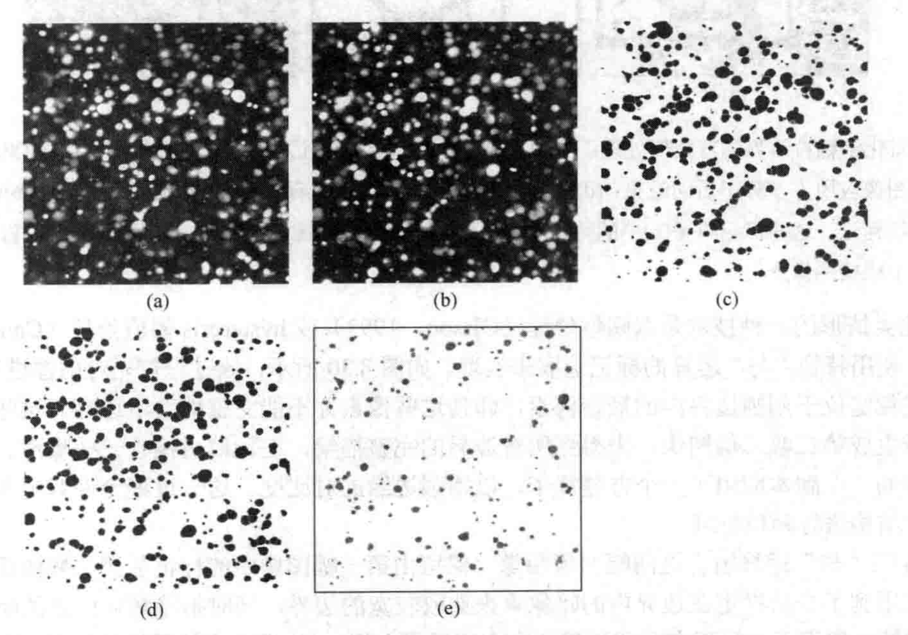


图 8.29　应用 Disector 逻辑：(a)和(b)共焦光学显微图像的两个平行截面；(c)图(a)经阈值处理后分隔了颗粒的结果；(d)图(b)经阈值处理后分隔了颗粒的结果；(e)使用标记选取只保留了图(c)和图(d)中的那些与另一幅图像不相交的特征，生成了将被计数的 200 多个特征（例子已用颜色显示了相交的平面）

下面的说明）。执行每个特征“与”运算，最终图像是“或”运算组合，它显示了在两个切片之一中出现的那些特征。

8.7 双阈值处理

对复杂图像进行阈值处理时，会用到另一种标记选择方法，如识别材料中的颗粒边界或组织中细胞的边界。在标本制备时，通常不能均匀地对细胞或颗粒的边界进行刻蚀或着色。图 8.30 所示例子是由颗粒内部的热蚀刻引起的。直接对图像进行阈值处理，并不能得到蚀刻边界的完整表示，更不用说颗粒内存在噪声的情形。在扫描印刷文本时，会出现一种相关的问题，即在纸张反面打印时可能会出现透印现象（Estrada & Tomasi，2009）。

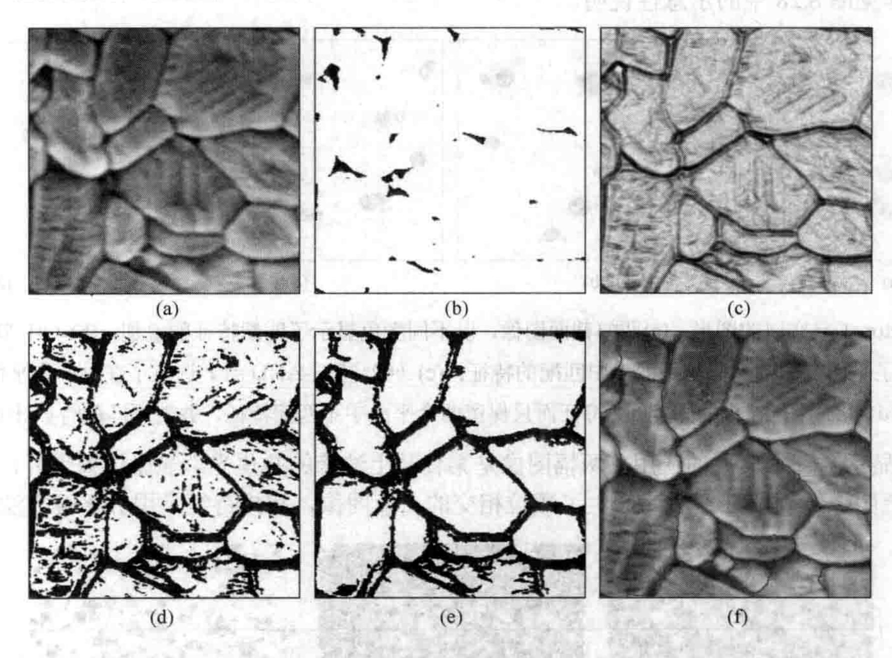


图 8.30　氧化铝颗粒边界的双阈值处理：(a)原始图像；(b)暗色颗粒边界标记经阈值处理后的结果；(c)对原始图像应用方差算子后的结果；(d)图(c)经阈值处理后得到了所有边界和其他标记；(e)从图(d)选择的特征标记，这些标记与图(b)中的特征相接；(f)图(e)经闭运算和骨架化处理后的结果，经修剪后叠加到了原始图像上

针对此类情形的一种技术是双阈值处理（Olsson，1993）或 hysteresis 阈值处理（Canny，1986）。这种技术由使用特征“与”运算的标记选取来实现。如图 8.30 所示，该过程首先对图像进行阈值处理，以选择那些肯定位于刻蚀边界内的最暗像素，即使这些像素并不能完整地表示边界。这些区域就成为标记。然后生成第二幅二值图像，来得到所有边界的完整描绘，在颗粒内接收一些噪声。在该例中，对原始图像的一个副本应用了一个方差算子，以增强边缘的对比度。这一过程允许对更多的边界及一些内部颗粒结构进行阈值处理。

使用特征“与”运算组合这两幅二值图像，保留由第一幅图像中的标记从第二幅图像中选取的任意特征。这用到了少量肯定在边界内的暗像素来选取较宽的边界，同时拒绝颗粒内部的噪声。为填充边界内的间隙，使用了一种形态学闭运算（在本章后面介绍）。边界宽度的明显增加并不重要，因为可使用骨架化（下面介绍）操作将边界线降低到最小宽度（实际的颗粒边界只有几个原子厚）。然后，可用修剪过的结果来测量颗粒及其边界。

在上例中，颗粒边界网络是图像的连续曲面细分。因此，可使用非双阈值方法的其他标准进行选择（如接触场中的对边）。图 8.31 显示了需要双阈值方法的一个例子。声学显微图像显示的是纤维加强材料的一个截面。这些图像本质上是带噪的，但双阈值处理（在本例中选取亮像素）允许选取围绕纤维的边界。因为纤维互相接触，因此还需要使用下面讨论的分水岭分割法将它们隔离出来以进行测量。

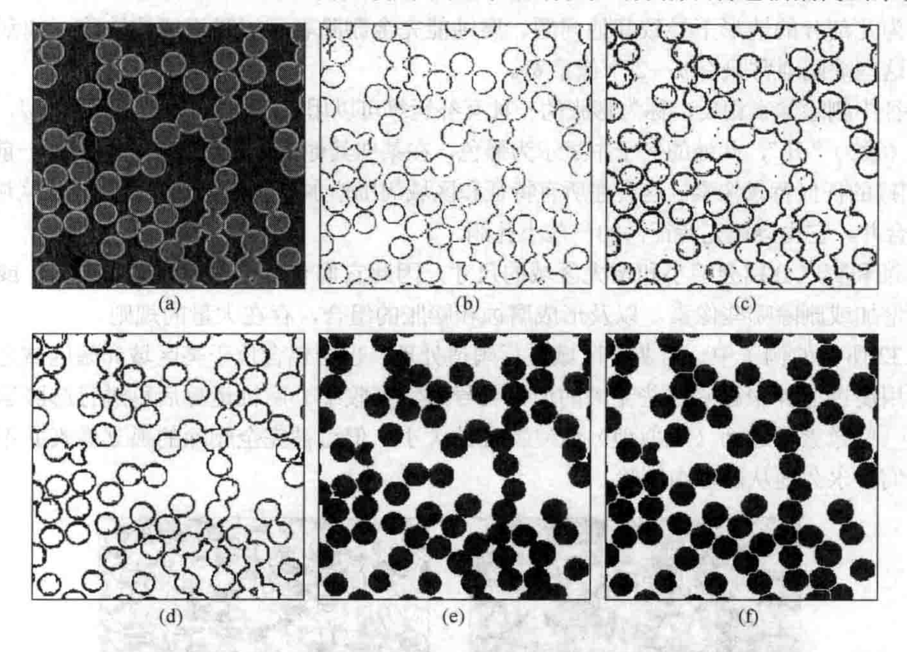


图 8.31　纤维边界的双阈值处理：(a)原始图像；(b)首次阈值处理后的结果；(c)第二次阈值处理后的结果；(d)标记选取结果；(e)填充后的边界；(f)分割后的纤维

8.8　腐蚀和膨胀

广泛用于处理二值图像的运算通常称为形态学过程（Serra，1982；Coster & Chermant，1985；Dougherty & Astola，1994，1999；Soille，1999；Shih，2009）。这些运算包括腐蚀和膨胀，以及这些运算的改进和组合。这些运算的经典形式基本上是邻域运算，如第 4 章和第 5 章中所示，其中使用了与邻域中像素值排序相类似的过程，来在空间域中处理灰度彩色图像。因为二值图像中的像素值限制为黑色或白色，因此操作更简单，通常只涉及计数而非排序。但基本思想相同，且在某些软件包中，排序法既用于连续色调灰度图像和彩色图像（使用像素亮度值），也用于二值图像。

数学形态学领域的相关文献很多。人们对这种运算开发了专用的语言和符号表示，且常以集合论来展现。此处采用一种更简单的方法：运算可以按照某种规则从二值图像中增加或者减少像素来描述，这些规则取决于邻域像素的模式。每个运算使用像素的原始模式在原始图像中的每个像素上执行。实际中，可能不需要创建一幅全新的图像，通过一次复制几行，可在内存中替找原有图像。在计算邻像素模式时，只用到原始像素值。

腐蚀操作会从图像内的特征中删除像素，方法是将原来的“开”的像素转换为“闭”像素。目的是删除那些不应出现的像素。最简单的例子是已经被阈值处理选择的像素，因为它们在感兴趣的亮度范围内，但不在具有该亮度的更大区域内。相反，它们也可能因为图像中的有限噪声而偶然具有那个亮度值，或者因为它们恰好跨越浅色和深色区域之间的边界，而恰好具有阈值处理所选择范围内的平均亮度。

这些像素无法通过简单的阈值处理来区分，因为它们的亮度值与期望区域的亮度值相同。使用布

尔逻辑删除它们是可能的，例如使用灰度级作为一个标准，使用梯度作为第二个标准，并要求将被保留的像素具有想要的灰度级和较小的梯度。但二值图像中通常仍然包括无关的像素。

最简单的腐蚀是删除（设为“闭”，此处的例子中显示为白色背景）与任一背景像素（已为“闭”）相接的任何前景像素。这会从所有特征和区域的周围删除一层像素，缩小特征或区域的尺寸，且会在使得特征分为几部分的情形下导致其他问题。腐蚀能完全删除表示点噪声或线缺陷（如刮痕）的无关像素，因为这些缺陷通常只有 1～2 个像素宽。

与从特征中删除像素相反，称为膨胀的一种互补运算可以用来添加像素。与腐蚀类似，经典的膨胀规则是添加（设为“开”，此处的例子中显示为黑色，在某些其他情形显示为彩色）与任一前景像素（已为“开”）相接的任何背景像素。这会在所有特征和区域周围增加一层像素，增大特征或区域的尺寸，且会导致特征合并。它也会填充特征内的一些小孔洞。

因为腐蚀和膨胀分别会缩小和增大区域的尺寸，因此它们有时也称为蚀刻和覆盖，或称为收缩和生长。决定增加或删除哪些像素，以及形成腐蚀和膨胀的组合，存在大量的规则。

在图 8.32 所示的例子中，对灰度区域进行阈值处理，也会包含位于亮区域和暗区域之间的一些灰色像素。使用腐蚀运算来删除这些无关的像素会导致特征收缩。腐蚀运算后再进行膨胀运算，会在特征外围增加一些像素，进而（近似地）恢复特征的大小。但已被完全删除的孤立像素并不会添加任何新像素，它们会永久地从图像中删除。

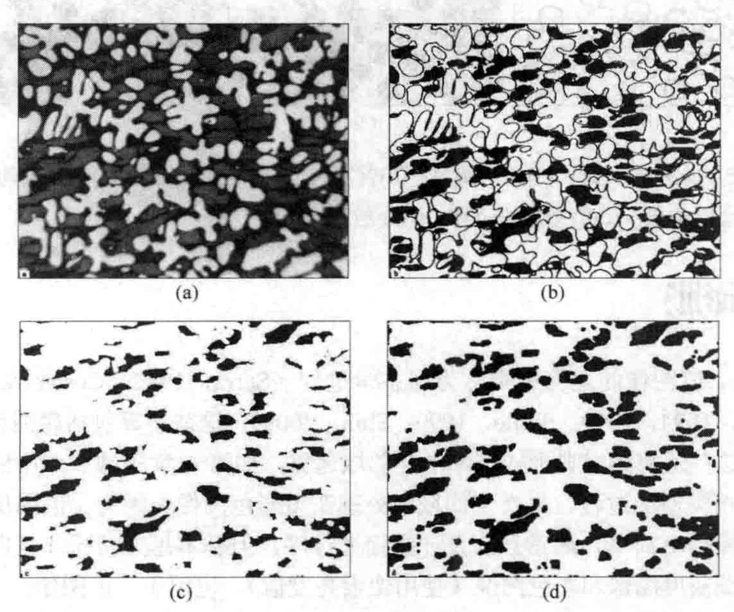


图 8.32 删除跨越边界的像素：(a)三相金属的原始显微图像；(b)中间灰度相经阈值处理后得到的二值图像；(c)使用两次迭代来腐蚀图(b)后的结果；(d)使用相同的两次迭代来膨胀图(c)的结果，注意恢复了特征的尺寸，但线条已不再存在

8.9 开运算和闭运算

先腐蚀再膨胀的组合称为开运算，指的是该组合打开恰好相接特征之间间隙的能力，如图 8.33 所示。这是在二值图像中移除模糊的线条和独立的噪声像素时一套很常用的操作。这是从二值图像中删除细线和孤立像素噪声的常用顺序组合之一。反序执行相同的运算（先膨胀再腐蚀）会产生不同的结果。这种顺序的组合称为闭运算，因为它会填充特征中的间隙。有一些参数可以用来调整腐蚀和膨胀

运算，特别是用来添加或删除像素的邻域模式或规则，以及迭代的次数，详见下面的讨论。在多数开运算和闭运算中，这些参数对腐蚀和膨胀而言相同。

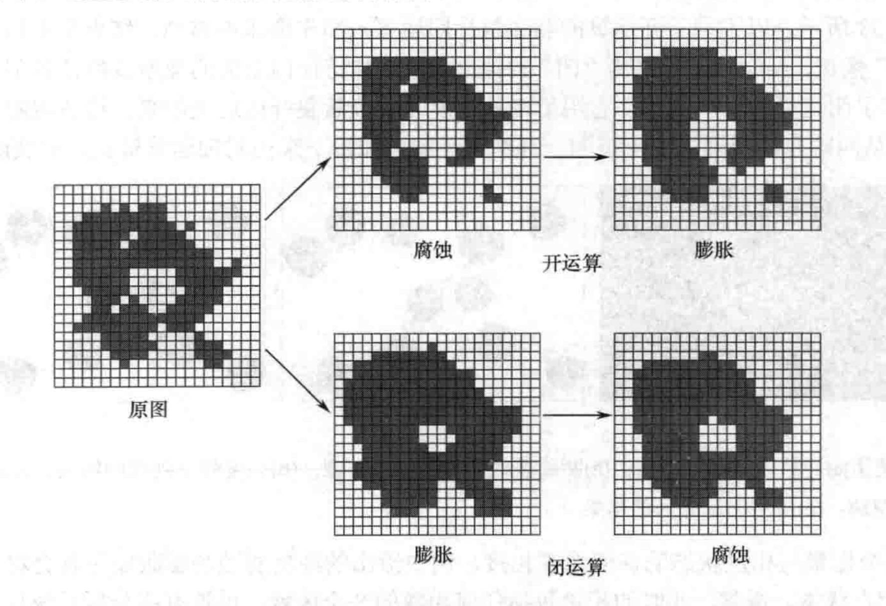


图8.33　组合腐蚀和膨胀来产生开运算或闭运算。应用两个运算的顺序不同，其结果也不同。由于原始图像是模糊的，因此有必要使用先验知识来选择正确的组合

在有些情形下，开运算可以分开相接的特征。在图 8.34 所示的例子中，所有特征的尺寸基本相同，彼此相接但重叠不多。这一事实使得我们可以继续腐蚀，直到所有特征完全分开而没有特征被完全删除。分离操作完成后，膨胀运算会将特征恢复为原始尺寸。它们可能会再次合并，除非使用某种逻辑来防止合并。如果某个像素的邻域属于不同的特征，那么阻止将该像素转换为“开”的一个规则会维护图中所示的分离。这种“非合并膨胀”需要对每个像素进行特征识别，所以在膨胀的每一步需要另外的逻辑。另一个规则阻止原始图像中任何不为“开”的像素转换为“开”，这样就将特征限制为它们的原始尺寸。

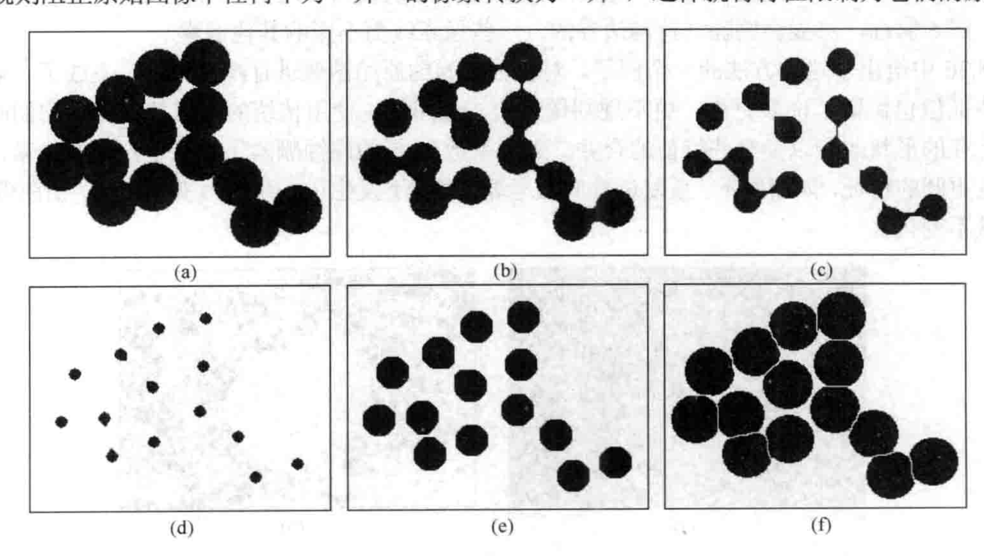


图 8.34　通过腐蚀/膨胀分开相接的特征：(a)原始测试图像；(b)两次腐蚀后的结果；(c)4 次腐蚀后的结果；(d)7 次腐蚀后的结果（所有特征均已分离）；(e)图(d)经 4 次膨胀后的结果；(f)几次非合并膨胀后，恢复了特征尺寸和边界

如果特征有不同的原始尺寸，那么该方法产生的分割线在相交处并不能正确地定位，在分开有些特征之前，其他特征可能会完全消失。下面介绍的分水岭分割技术可以更好地应付这种情况。

如图 8.33 所示，以不同于开运算的顺序执行闭运算，即先膨胀再腐蚀，结果是不同的。替代删除孤立的“开”像素，结果填充那些为“闭”的孤立像素（如特征内丢失的像素或特征各部分间的间隙）。图 8.35 显示了闭运算的一个例子，它用于将截面中的部分断裂纤维连接起来。这些裂隙都很窄，因此膨胀运算会从两端延伸像素来连上裂隙。纤维直径的增加接下来由腐蚀运算修正，但裂隙不再出现。

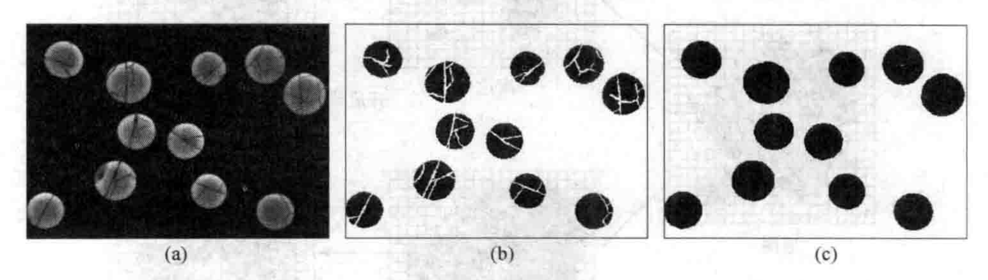


图 8.35　使用闭运算连接部分特征：(a)断裂玻璃纤维的原始图像；(b)亮度阈值处理的结果，显示了光纤内的裂隙；(c)应用闭运算后的结果

如果某个像素与相反状态的任何像素相接，例中给出的经典腐蚀和膨胀操作就会将该像素转换为“开”或“闭”状态。通常，此时的相接包括任何相邻的 8 个像素，尽管有些分析系统只针对 4 个共享边缘的邻域。对于六边形像素阵列，这些运算将会更加简单，并更具有各向同性，因为像素邻域距离都是一样的，但从实用考虑使得我们更常使用方形像素网格。

还存在大量的其他规则。一种方法是条件腐蚀或膨胀，它统计有着相反颜色或状态的邻域像素的数量，并将该值与一个阈值进行比较，当该测试系数超过阈值时才改变中间像素的状态。在这种方法中，经典的腐蚀对应于零系数。不同系数值的效果之一是改变特征的增长率或收缩率，并在某种程度上控制结果的各向同性。这将在下面说明。

也可选择一个较大的系数（5～7），来只选择那些孤立的噪声像素而保持多数特征不变。例如，系数为 7 时只有各个孤立的像素被反转（在腐蚀运算中删除或设为“闭”，在膨胀运算中则与此相反）。以系数 5 或 6 腐蚀，可能会删除（跨越边界的）一些像素线而不影响其他像素。

图 8.36 中给出了这种方法的一个例子。对色素细胞的原始图像进行阈值处理，生成了一幅显示出感兴趣特征但也保留了许多更小、更不规则像素的二值图像。使用传统的开运算来删除它们时，会导致较大特征的形状改变以及有些特征的合并。邻域系数为 5 的腐蚀删除了小且不规则的像素，但未影响较大且更圆的特征，如图所示。重复进行腐蚀运算，直到不发生进一步的改变为止（二值图像中“开”像素数量不变化）。

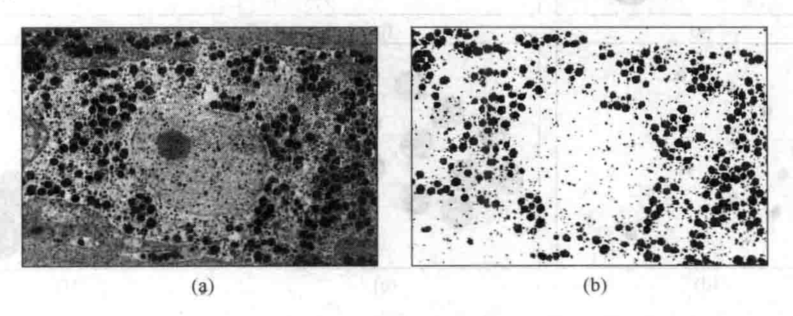
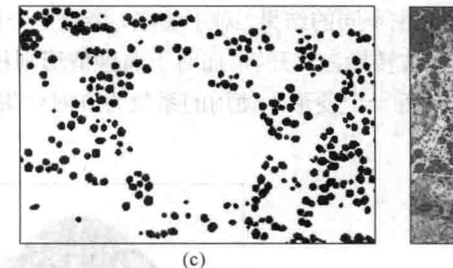


图 8.36　删除图像中的碎片：(a)染色细胞的原始图像；(b)亮度阈值处理结果，显示了染色颗粒和其他小而不规则的特征；(c)开运算（邻域系数为 5）保留了大而规则的颗粒；(d)叠加到原始图像上的颜色表明，红色是腐蚀后保留的特征，蓝色是已被删除的特征



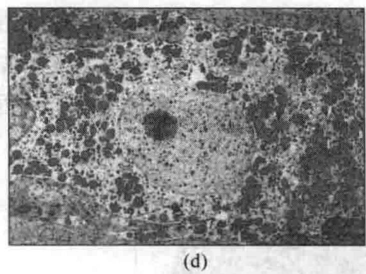


图 8.36 删除图像中的碎片：(a)染色细胞的原始图像；(b)亮度阈值处理结果，显示了染色颗粒和其他小而不规则的特征；(c)开运算（邻域系数为 5）保留了大而规则的颗粒；(d)叠加到原始图像上的颜色表明，红色是腐蚀后保留的特征，蓝色是已被删除的特征（续）

图 8.37 中的测试图像中显示了许多细线和窄隙，它们可用不同的邻域系数和迭代次数（先腐蚀再膨胀的次数，或先膨胀再腐蚀的次数）来删除或填充。

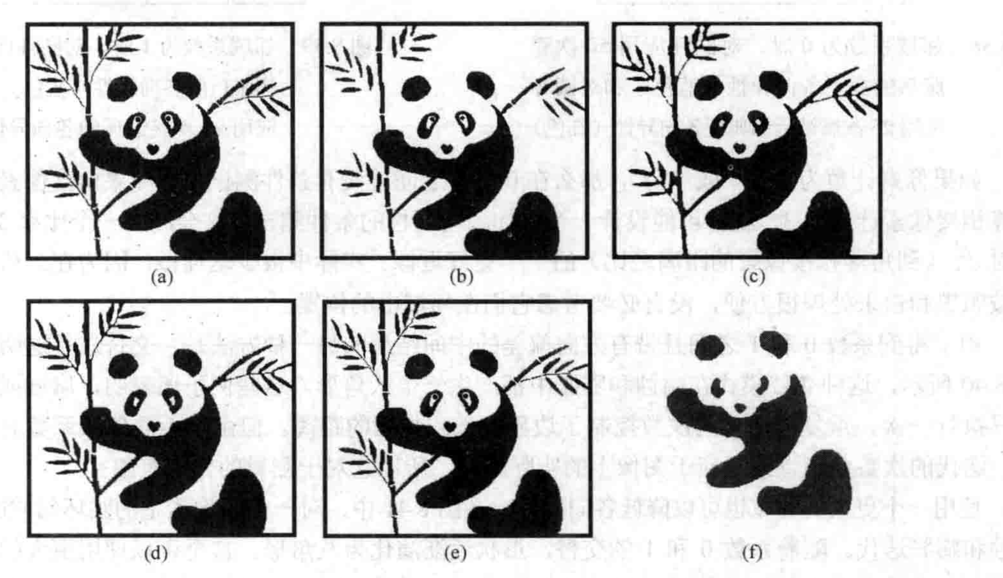


图 8.37 不同邻域系数和迭代次数的效果说明：(a)原始测试图像；(b)邻域系数为 3 的一次腐蚀，删除了孤立的线条和点；(c)邻域系数为 2 的 2 次闭运算，填充间隙使得特征相连的同时，删除了孤立的点；(d)经典的闭运算（邻域系数为 0，1 次迭代）连接了多数特征，但保留了孤立的点；(e)邻域系数为 7 的一次开运算，删除了噪声点而未影响其他内容；(f)邻域系数为 1 的 4 次开运算，删除了包含图框在内的所有小特征

8.10 各向同性

不可能对一个 3×3 的小邻域来定义各向同性的邻域模式。适用于圆环的经典腐蚀并不能均匀地收缩圆环，而在 45° 对角线方向会以更快的速率进行，因为在这些方向上的像素间距更大。其结果是，圆环腐蚀会趋向于菱形，如图 8.38 所示。一旦某个特征达到这个形状，它就会继续均匀地腐蚀，并保持该形状。但在多数情形下，特征并不是真正的菱形，而是表现为一种严重的失真。

同样，适用于圆环的经典膨胀在 45° 对角线方向也会进行得更快，因此膨胀趋向于正方形，如图 8.38 所示。方形在膨胀中是稳定的，但膨胀中出现的块状图像失真，会影响到下一步分析。

邻域系数为 1 而非 0 的条件腐蚀和膨胀会产生不同的结果。对于膨胀，接触 1 个以上前景像素（即 8 个可能邻域位置中的两个或更多）的背景像素被转换为“开”，而对于腐蚀情形则相反。使用这个过程腐蚀圆环趋向于一个正方形，而膨胀趋向于接近一个菱形，使用的系数为 0 时结果相反。图 8.39 中显示了这一结果。

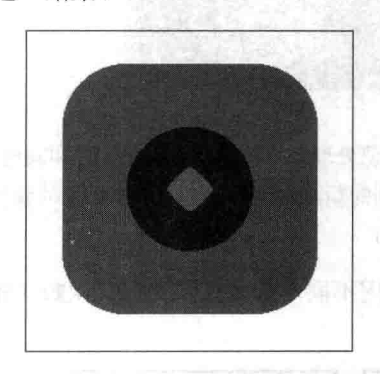


图 8.38　邻域系数为 0 时，对圆环应用 50 次膨胀后的典型各向异性（蓝色）和对圆环应用 25 次腐蚀后的典型各向异性（红色）

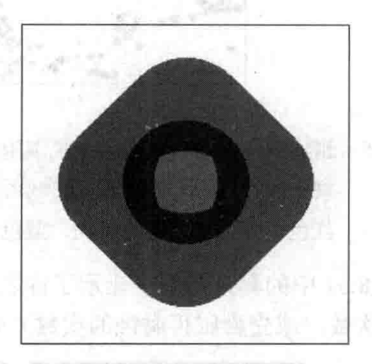


图 8.39　邻域系数为 1 时，对圆环应用 50 次膨胀后的各向异性（蓝色）和对圆环应用 25 次腐蚀后的各向异性（红色）

如果像素计数为“开”或“闭”，那么在 0 与 1 之间就没有条件测试值。如果角落像素计数为 2 而边缘相接像素计数为 3，就有可能设计一个更加各向同性的条件测试。这会产生一个比率 3/2 = 1.5，它是对 $\sqrt{2}$（到角落相接像素的距离之比）的一个更好近似。实际中很少这样做，因为在二值图像中将像素按照黑和白来处理很方便，没有必要考虑它们在邻域中的位置。

用于得到系数 0 和 1 之间且带有方向偏差的中间结果的另一种方法是，交替进行这两个测试。如图 8.40 所示，这种交替模式在腐蚀和膨胀中都产生一个八角形。这些例子还表明，腐蚀或膨胀的点不必只执行一次。重复或迭代的次数控制了边界增长或收缩的距离。但由于不同邻域系数对该过程的影响，迭代的次数并不直接对应于图像上的实际距离，所以它对于测量的作用有限。

使用一个更大的邻域也可以降低各向异性。在图 8.41 中，对一个 5 像素宽的圆环邻域进行了 10 次腐蚀和膨胀迭代。随着系数 0 和 1 的交替，形状逐渐演化为八角形。甚至可以使用更大的邻域，但这也增加了需要对各个邻域计数的时间，且在迭代不产生各向同性的腐蚀和膨胀时，对受腐蚀和膨胀影响的距离的控制更弱。从原理上来讲，经典的腐蚀或膨胀（系数为 0）以及一个半径大到足以允许单次应用而非迭代的圆形邻域，可以产生各向同性的结果，但实际情形并非总是如此。

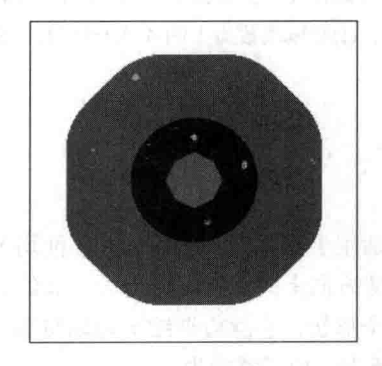


图 8.40　交替使用邻域系数 0 和 1，对圆环应用 50 次膨胀后的各向异性（蓝色）和对圆环应用 25 次腐蚀后的各向异性（红色）

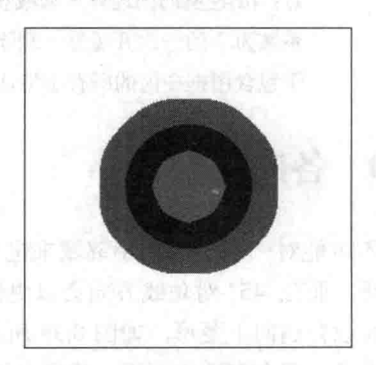


图 8.41　使用 10 次迭代和一个更大的邻域（近似于圆的 5 像素宽度）的经典腐蚀（红色）和膨胀（蓝色）

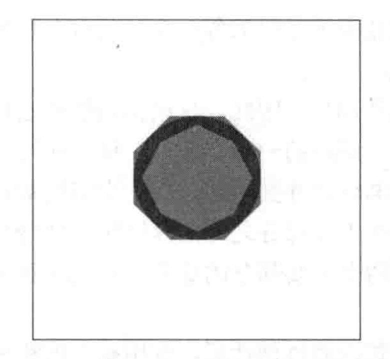


图 8.42　使用邻域系数 3，对圆环应用 50 次膨胀后得到的八边形（蓝色，最初几次迭代后无变化）和对圆环应用 25 次腐蚀后得到的八边形（红色），注意添加和删除的速率较低

每个邻域模式或系数都有其典型的各向异性。图 8.42 显示了邻域系数为 3 时的有趣结果。和交替应用 0 和 1 模式类似，该运算产生了一个八边形。但腐蚀率很低，且膨胀使得该图形增长到了外接八角形，然后趋于稳定而不再进一步增加像素。这个系数有时会用于构建围绕特征的外接多边形。

8.11　使用腐蚀和膨胀进行测量

使用系数 0 或 1，或交替使用它们，进行 *n* 次腐蚀，会使得特征径向地收缩 *n* 个像素（局部变化取决于原始特征的形状）。这会使得尺寸小于 2*n* 个像素的特征完全消失。统计在每一次迭代中消失的特征的数量，就可得出远小于该尺寸的特征数量。这意味着腐蚀和计数可用于了解尺寸的分布情况，而不用实际测量特征（Ehrlich et al., 1984）。

图 8.43 说明了这一方法。从特征边界外部删除的连续像素层以彩色显示。统计每次迭代后消失的特征数量，就得到了所示的图形，该图形近似对应于相接且稍有重叠的圆环的尺寸分布。为便于比较，对图 8.57 应用分水岭处理（见下面的介绍），得到了分离的特征，且测量这些特征的尺寸产生了如图所示的更为精确的结果（对于一个外接圆的直径），进而可根据大小来区分各个圆。

图 8.43　迭代腐蚀和消失特征的数量统计：(a)一组外接且稍有叠加的圆；(b)连续迭代删除的像素层，已表示为不同的颜色；(c)每次迭代消失的特征数量；(d)分水岭分割后，各圆的尺寸分布（如后面的图 8.57 所示）

使用灰度形态学（即第 5 章中介绍的一种排序运算，它使用较亮或较暗的邻域来替代每个特征）并统计消失的特征的数量，可采用同样的方法。这将在下面介绍。

对于不规则形状和凹状特征，腐蚀过程可能会使得特征分成几部分。因此，仅根据腐蚀的迭代次数与特征数量的关系来了解尺寸分布并不是一种好方法。解决这一问题的一种方法是，使用相同的系数和步数腐蚀后，再进行膨胀。这会把许多（但非全部）分开的部分合并到一起，并给出特征数量的更好估计。但原始特征的形状仍具有相当大的敏感性。在处理两个主要部分之间的腐蚀时，哑铃形的物体会被分成两部分，它们并未合并在一起。如果目的是统计这两个主要部分的数量，那么这种分离可能是需要的。

也可以使用上面介绍的特征“与”运算或标记选择方法。在每次腐蚀迭代后，使用剩下的特征来只选择那些与它们相接的原始特征。对原始特征进行计数可给出正确的数量。在功能上这等同于在图像中的每个像素上保留特征标签，并统计每次腐蚀后图像中出现的不同标签的数量。这种使用任何一种校正技术来估计尺寸分布而不测量特征的方法，尤其适用于地质测量，如测量矿物颗粒或沉淀物的大小。

执行膨胀并根据特征与步数的关系来统计单独特征的数量的相反方法并不常见。这种方法可给出图像中各特征之间最近距离的估计。采用传统的特征测量方法时，先确定每个特征的 x, y 位置，然后对所生成的数据文件排序，进而求出最近的邻域及其距离。此时，距离是指从中心到中心的距离。当特征明显大于它们间的间距，或特征的形状很重要时，表征它们的边界间的距离（边缘到边缘的距离）会更重要可能会更有趣。膨胀方法可以提供这种信息。

当计算机的能力有限且更精确的方法不切实际时，可使用前几段中介绍的方法，尽管这种方法存在各向异性问题，但这些方法大部分已被使用欧氏距离图的方法所取代，如下所示。

替代统计每次腐蚀迭代后消失的特征的数量，统计仍然为“开”的剩余像素的数量更为容易，因为它可提供边界形状的一些信息。平滑的欧几里得边界以固定速率腐蚀。不规则边界和分形边界则不以固定速率腐蚀，因为有更多的像素位于边界上并与对面的邻域相接。尽管存在更精确的方法，但这种效果已用于估计分形维数，讨论如下。

分形维数以及基于自相似粗糙度将边界描述为分形，是相当新的思想，这些思想可在科学和艺术的许多应用中找到（Mandelbrot，1982；Feder，1988；Russ，1994）。此处并不介绍这些思想的有趣背景和概念的使用。通过腐蚀和膨胀来测量分形维数的基本思想源自分形边界维数的闵可夫斯基定义。将膨胀一个区域的结果与（腐蚀该区域得到的）另一幅图像进行“异或”运算，可得到沿边界的像素。

为测量分形维数，可通过逐渐增加腐蚀和膨胀的迭代次数来重复该运算（Flook，1978），并画出边界的有效宽度（像素总数除以长度和迭代的腐蚀/膨胀周期数）与迭代次数的对数关系图。对于欧几里得边界，该图显示不出趋势；“异或”运算沿边界选取的像素的数量，随腐蚀/膨胀的迭代次数线性增加（即每次迭代影响同样数量的像素）。然而，对于具有自相似细节的粗糙边界，图中对数轴上显示了线性变化，变化的斜率直接给出了边界的分形维数，如图 8.44 所示。

还有许多确定边界分形维数的其他方法，包括计盒法或镶嵌融合法［这种方法随着图像分辨率变粗所致的像素尺寸的增大，统计边界穿过的像素数量（Kaye，1986；Russ，1990）］和等步进法［该方法需要将边界表示为多边形而非像素（Schwarz & Exner，1980）］。对于分形边界，这些方法也会生成对数直线图，由直线的斜率可求出分形维数。第 11 章中将介绍执行这种测量的几种更精确的技术。

根据特征与膨胀的关系来统计像素数量，可间接测量特征聚类，因为随着附近特征的合并，边界数量会减少，且区域的生长速率会降低。只统计像素而不统计特征很难区分边界形状和特征间距的影响。如果所有特征最初都非常小，或如果它们由单个点（如重心）表示，那么这种方法可为聚类提供分形维数（技术上称为 Sierpinski 分形）。

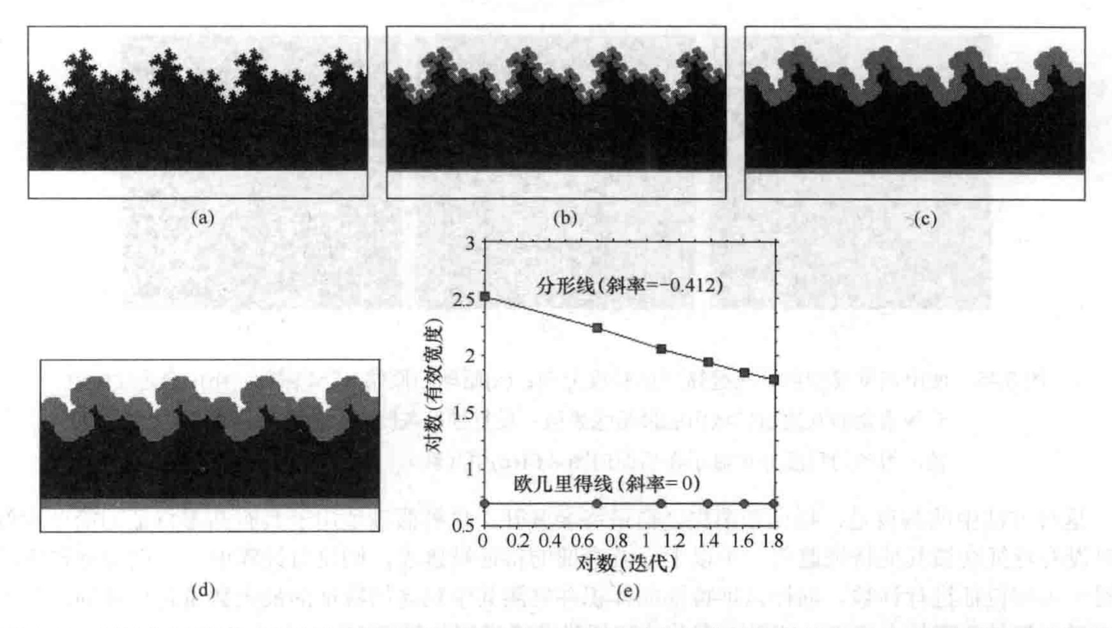


图 8.44　通过腐蚀/膨胀来测量闵可夫斯基分形维数：(a)顶部边界为典型的科赫分形而底部边界为欧氏直线的测试图像；(b)灰色像素显示了 1 次腐蚀和 1 次膨胀结果间的不同；(c)2 次腐蚀和 2 次膨胀结果间的不同；(d) 4 次腐蚀和 4 次膨胀结果间的不同；(e)有效宽度（灰色像素的面积除以图像的宽度和迭代次数）与迭代次数（约为灰色带的宽度）间的对数关系图。分形维数等于 1 减去斜率 1.412

8.12　扩展到灰度图像

第 4 章和第 5 章中介绍的一种图像处理运算是排序算子，它在邻域中找到最亮或最暗的像素，并用它来替换中心像素。这一运算称为灰度腐蚀或膨胀，具体取决于是使用最亮像素值还是使用最暗像素值来使得特征增长或收缩。对灰度和彩色图像应用形态学运算类似于对二值图像的形态学运算，如果将灰度形态学程序应用到二值图像，那么它们就相当于经典的腐蚀/膨胀运算。

正如腐蚀二值图像中的特征可得到特征尺寸的分布那样，也可对连续色调图像进行灰度腐蚀。图 8.45 给出了一个例子。这幅 SEM 图像中的脂球局部堆积在一起并相互接触，因此很难使用传统的图像测量技术来处理它。应用灰度腐蚀会减小特征尺寸，统计每步腐蚀后消失的中心亮点的数量，可得到尺寸分布［为与一种更好的方法比较，这种尺寸分布显示在下面的图 8.61(d)中］。

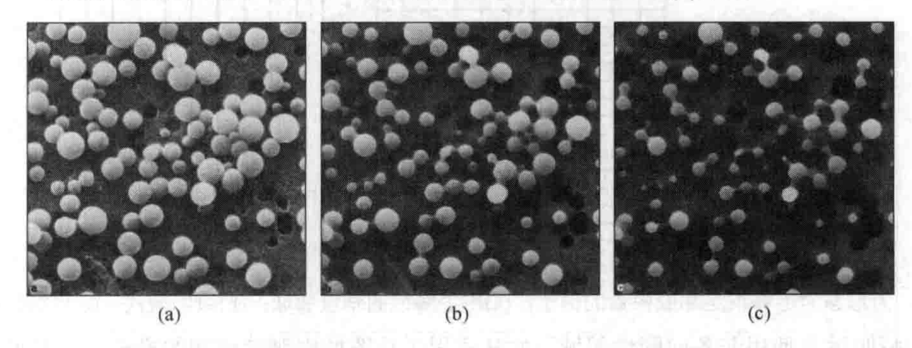
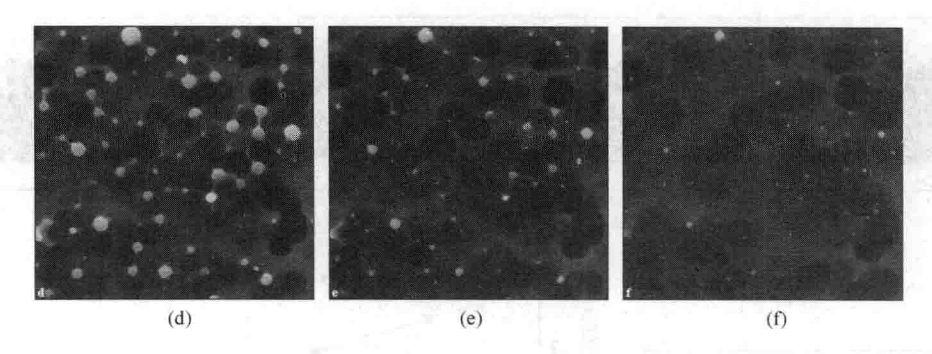


图 8.45　使用灰度腐蚀估计重叠微粒的粒度分布：(a)脂球的原始 SEM 图像；(b)~(f)通过保留 5 像素宽的八边形邻域内的最暗像素值，反复进行灰度腐蚀的结果。为与其他方法比较，得到的粒度分布显示在后面的图 8.61(d)中



(d) (e) (f)

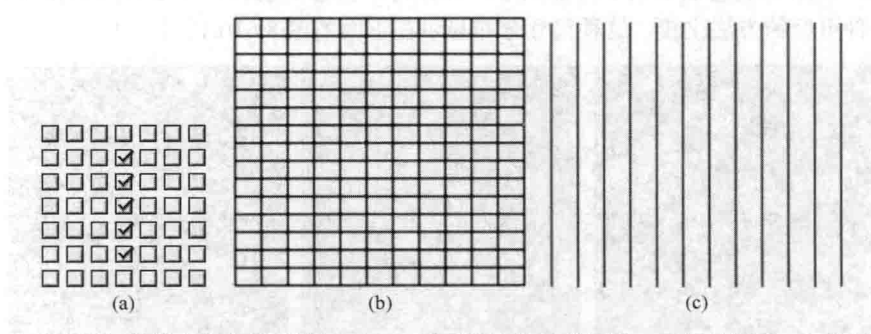
图 8.45 使用灰度腐蚀估计重叠微粒的粒度分布：(a)脂球的原始 SEM 图像；(b)~(f)通过保留 5 像素宽的八边形邻域内的最暗像素值，反复进行灰度腐蚀的结果。为与其他方法比较，得到的粒度分布显示在后面的图 8.61(d)中（续）

这种方法中的假设是，特征在消失之前最终会分开。这种假设适用于具有凸形特征的简单图像，此时没有特征会被其他特征遮盖一半以上。若内部的特征被遮盖，则没有纯粹的二维图像处理方法可以对一大堆特征进行计数。估计该堆特征的体积并猜测其中包含的特征的最大数量是可能的，但不可能知道它们是否真的在那里，或是否有其他特征隐藏可见层的最下方。

8.13 形态学邻域参数

特定邻域模式也可以用于腐蚀和膨胀运算。最常见的模式是，将中心像素与其 4 个边缘相接的邻域进行比较（由于邻域形状的原因，这种模式通常称为“+”模式），或将中心像素与其 4 个角点相接的邻域进行比较（这种模式称为“x”模式），如果这 4 个邻域具有相反的颜色，就改变中心像素。它们很少单独使用，但可用在交替序列中，以得到比经典腐蚀更大的方向一致性，类似于交替使用系数 0 和 1 对测试产生的效果。

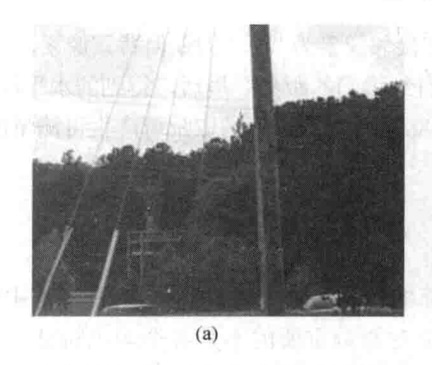
如果某种特定的邻域模式适用于特定的应用，则可使用它们。甚至不需要将比较限制到直接接触的邻域。至于灰度运算，更大的邻域可使得它对应于更细的纹理，并对方向性实现更大的控制。图 8.46 给出了一个简单的例子。这类运算的一般情况称为击中击不中算子，它指定一个分为三类的相邻像素模式：第一类是“开”像素，第二类是“闭”像素，第三类是可以忽略的无关像素。如果找到了一种模式，则将像素设置为指定的状态（Serra，1982；Coster & Chermant，1985）。



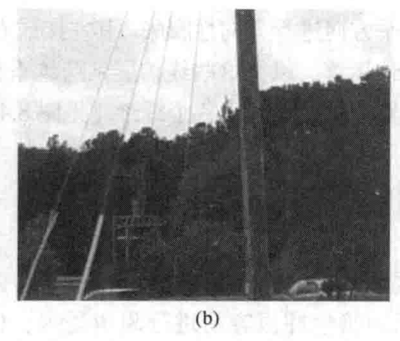
(a) (b) (c)

图 8.46 为形态学运算指定邻域像素的例子：(a)用于腐蚀的垂直邻域；(b)原始模式；(c)腐蚀后的结果

对腐蚀/膨胀运算使用非各向同性邻域，尤其适用于从图像中删除已知的模式。最常见的例子是从扫描底片中删除划痕。胶卷上的划痕通常很窄且是纵向的。比划痕宽度更长且与划痕垂直的邻域，可使用相邻图像中的值来替换划痕中的暗色像素。该方法也可用于删除其他的线性结构，如图 8.47 所示。该例已使用第 4 章中介绍的对灰度或彩色图像排序的方法执行了开运算（腐蚀并膨胀）。



(a)



(b)

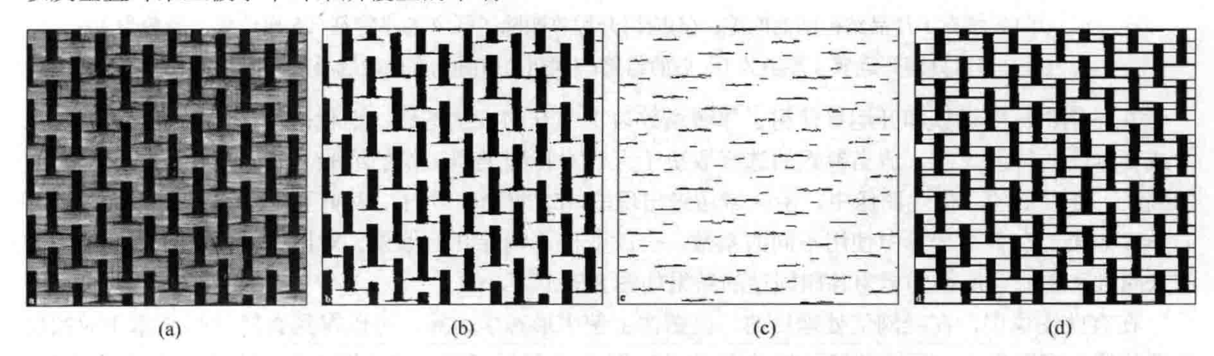
图 8.47 删除电线：(a)原始图像；(b)使用垂直邻域对图像进行开运算和闭运算后的结果。注意只删除了水平线，删除其他线时需要不同的邻域模式

这种类型的运算通常称为模板匹配，或针对灰度图像时称为相关（如第 6 章中使用傅里叶变换实现的图像），它是搜索图像中特定模式的一种方法。对二值图像，它同样适用；模板与阈值处理后的二值图像匹配，是针对字符光学识别的最早方法之一，且仍然用于字符形状、大小和位置严格受控的情形下（如银行支票底部的字符）。读取更普通的文本则需要更灵活的方法。

命运表是一种可快速实现邻域比较的方法。这种方法认为 8 个最近邻像素的值为 1 或 0，具体取决于像素是“开”还是“闭”。将这 8 个值放入单独的位中产生一个字节，它可以有 256 个可能的值。然后使用该值作为可提供结果的一个表的地址（即将中心像素变为“开”或“闭”）。构造地址的有效方式使得该方法非常快。创建可能命运的几个表（可保存到磁盘上并在需要时载入）来处理不同腐蚀和膨胀规则的能力，使得该方法非常灵活。但它不能很好地推广到更大的邻域或三维体素阵列图像，因为此时表格会变得相当大。

应用不对称或各向异性的特定腐蚀/膨胀操作时，总是要求具有能从图像中选择性地提取的一些独立知识、期望的信息，以及选取的运算。但这并不像要求的那样重要，因为所有图像处理在一定程度上都是知识引导的。人类观察者会试图找到运算方法来提取因某些原因已知或希望出现的信息。

图 8.48 给出了一个例子。水平纺织纤维的宽度会随其织在垂直纤维上方和下方而变化。为对编织的机械性质建模，测量这种变化是非常重要的。暗色垂直纤维可以基于亮度进行阈值处理，但要勾画出水平纤维则比较困难。图中所示的过程使用了结构的已知方向性。暗色纤维经阈值处理后，执行一次腐蚀，以便仅删除那些其上下邻点刚好为背景一部分的像素。然后，将这些像素［如图 8.48(c)所示］与原始二值图像进行“异或”运算，可分离出这些像素。这些像素包括水平纤维间的少量可分辨点，以及垂直纤维上被水平纤维所覆盖的末端。



(a) (b) (c) (d)

图 8.48 使用方向性腐蚀和膨胀来分割一幅图像：(a)织物的原始灰度图像；(b)图(a)经亮度阈值处理后的结果；(c)执行垂直腐蚀并与原始图像“异或”运算后，隔离的末端像素；(d)重复水平膨胀图(c)并与原始图像“或”运算得到的完整结果

接着，在水平方向进行方向性膨胀。其左接或右接邻像素为“开”的任何背景像素，自身设为“开”，且重复足够多次该运算，以将这些线扩展到跨越垂直纤维间的距离。最后，得到的水平线与原始二值图像进行“或”运算，以便勾画出各个纤维［见图 8.48(d)］。对这幅图像取反可产生可测量的特征。

8.14　应用示例

定义测量区域时，通常须用到闭运算和开运算的组合。在图 8.49 中，钢件标本中的碳化铁粒子经蚀刻后，显示为暗色并很容易进行阈值处理。但对测量重要的不是各个暗色特征。片晶结构和没有这一结构的孤岛大小控制着金属的性质，但它们都不是由各个暗色的碳化颗粒定义的。先膨胀再腐蚀（闭运算）后，各个片晶结构合并到了一起，但因为原始图像中存在一些暗点，较大的白色颗粒内也存在暗色特征。先闭运算再开运算（整个顺序是膨胀、腐蚀、腐蚀、膨胀）生成了有用的结果，如图所示。

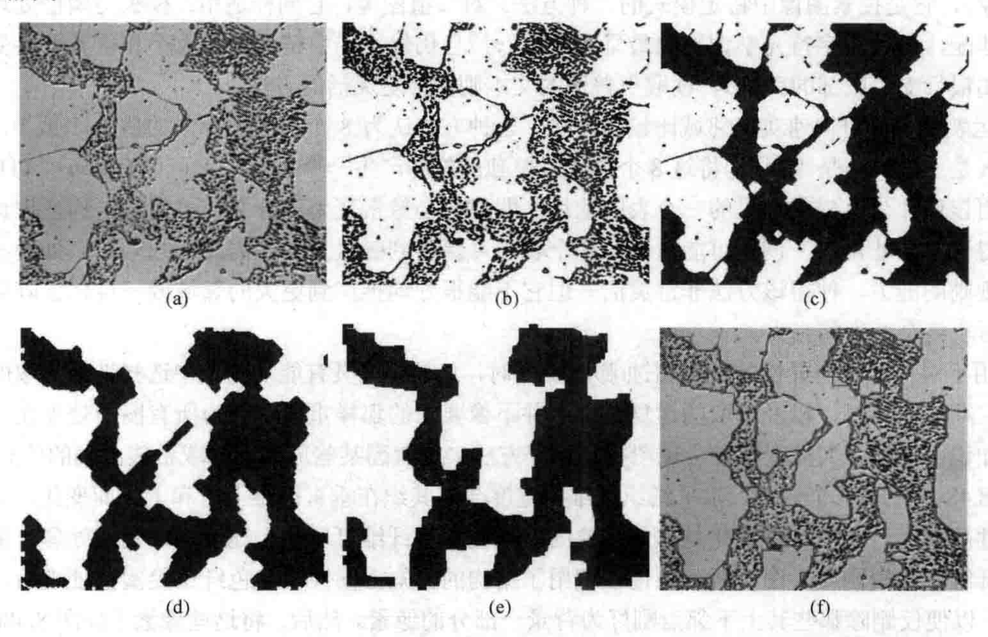


图 8.49　组合闭运算和开运算来描绘区域：(a)化学蚀刻的金相钢件标本的原始灰度图像（暗色区域是碳化铁，明亮区域是铁）；(b)对图(a)进行亮度阈值处理的结果；(c)应用闭运算（6 次迭代，系数为 1）填充了片晶结构间的间隙；(d)应用开运算删除了孤立的小特征（6 次迭代，系数为 1）；(e)先闭运算再开运算（系数为 0）后的结果；(f)图(d)和图(e)中的轮廓叠加到原始图像上的结果

在该例中，闭运算和开运算使用了邻域系数为 1 或 0 的 6 次迭代。迭代次数基于要填充的间隙尺寸或要删除的特征尺寸。适当参数的选择取决于人们对特殊类型图像的经验，以及对最终结果正确性的判断。处理后的二值化图像中，沿区域边缘出现的 45° 或 90° 方向，表明了腐蚀/膨胀运算导致的各向异性效果。在各个运算中使用不同的系数，会改变但不消除几何偏差。对同一幅图像使用不同的方法来腐蚀和膨胀，所得的更为各向同性的结果如图 8.54 所示。

在有些图像中，在经阈值处理后的二值图像上使用形态学运算，可以得到类似于第 5 章中对阈值处理前的灰度图像上应用纹理算子的结果。例如，图 8.50 显示了第 5 章中图 5.43 所用的相同凝乳图像。经背景调匀并对平滑的白色区域（凝乳）进行阈值处理，产生了所示的结果。图像内有纹理的部分中有许多区域，它们的亮度与凝胶的亮度相同。在灰度纹理处理时，可基于像素亮度的局部变化来删除

这些区域。在这幅图形中，该变化产生了狭窄且不规则的区域。由用于删除边缘相接像素的腐蚀和将像素还原到仍出现的特征边界的膨胀所组成的开运算，有效地删除了背景的杂乱，如图中所示。

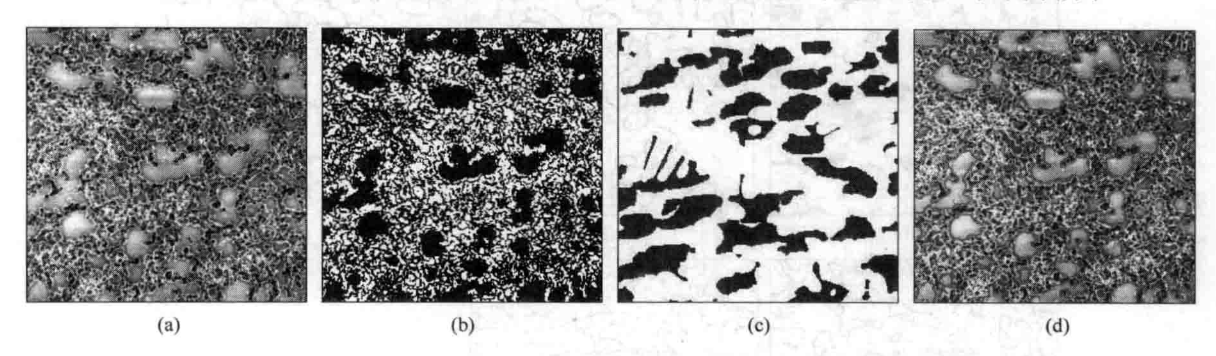


图 8.50　通过腐蚀/膨胀分割一幅图像：(a)原始图像（奶酪的凝乳）；(b)阈值处理后的结果；(c)形态学开运算后的结果；(d)叠加到原始图像上的轮廓

在该图像中定义结构的腐蚀/膨胀方法，意味着要对特征的尺寸、特征边界的不规则性和特征间的间距做一些假设，而第 5 章中成功的灰度处理要求的则是关于纹理空间尺度的一个类似决策。但通常而言，应在阈值处理之前对图像进行处理，以充分利用包含在像素亮度值中的信息，而不应在阈值处理后再处理图像。

腐蚀/膨胀过程通常与布尔逻辑组合使用。在图 8.24 和图 8.26 的例子中，用于测试邻接性的线条，是通过膨胀二值图像，然后将结果与原始图像"异或"产生的。图 8.49(f)和图 8.50(d)中所示的叠加轮廓，以及用来显示原始图像处理结果的其他轮廓，是通过先腐蚀再与膨胀后的二值图像"异或"产生的。这保留了原本与背景相邻的像素的轮廓，它可以作为一个颜色叠加。通过一种改进的腐蚀运算也可直接获得一幅轮廓图像，这种运算会删除没有相邻背景像素的任何像素。但它会产生称为卡斯特特征的轮廓，这一名称或许源自 1876 年曾被包围的乔治·阿姆斯特朗·卡斯特[①]。

卡斯特可用于确定特征或区域间的相邻关系。作为一个例子，图 8.51 显示了三相金属合金的光学显微图像。每个相都可通过阈值处理来描绘（且在中等灰度图像情形下，应用开运算删除了跨越黑白边界的几行像素，如图 8.32 所示）。然后可形成每相的卡斯特，如上所述。

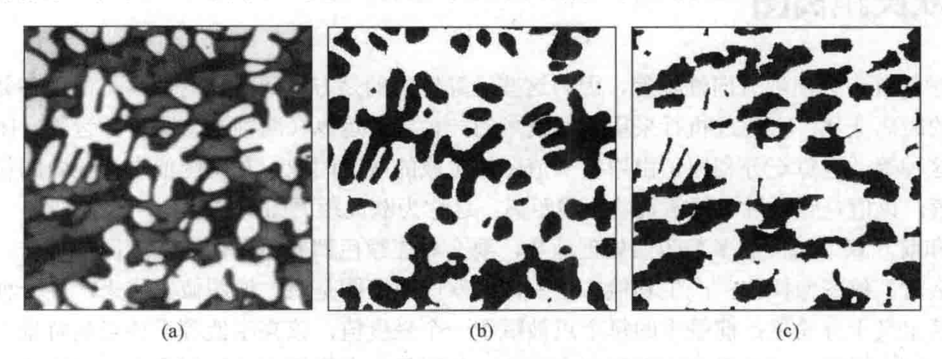


图 8.51　使用布尔逻辑来测量邻接关系：(a)三相金属的原始显微图像；(b)经阈值处理的白相；(c)经阈值处理的灰相；(d)经阈值处理的黑相；(e)白相的环绕轮廓；(f)灰相的环绕轮廓；(g)黑相的环绕轮廓；(h)以不同颜色组合的相对"与"轮廓，用来识别每相/相界面

① 乔治·阿姆斯特朗·卡斯特（George Armstrong Custer，1839—1876），美国骑兵军官，美国内战时联邦军将领。在 1876 年 6 月 25 日的"小毕霍恩河战役"中，印第安联盟酋长"坐牛"成功地把卡斯特引进了圈套，成功集结了比卡斯特的部队多出 4 倍的印第安人（印第安联盟 3000 多人，而卡斯特有 700 骑兵），全歼了卡斯特的军队并击毙了他。——译者注

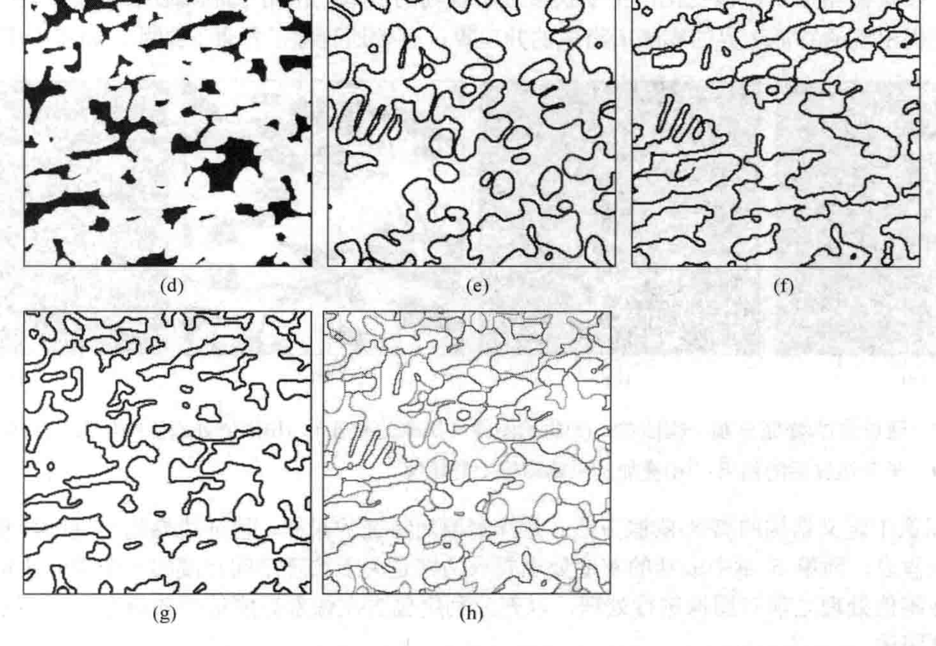


图 8.51　使用布尔逻辑来测量邻接关系：(a)三相金属的原始显微图像；(b)经阈值处理的白相；(c)经阈值处理的灰相；(d)经阈值处理的黑相；(e)白相的环绕轮廓；(f)灰相的环绕轮廓；(g)黑相的环绕轮廓；(h)以不同颜色组合的相对“与”轮廓，用来识别每相/相界面（续）

使用“与”运算组合每个相和其他相的卡斯特轮廓，令保留了两个相共有的部分卡斯特轮廓。其结果是将边界标记为白－灰色、灰－黑色或黑－白色，因此可确定每种类型的范围。第 9 章中将介绍如何使用剖面图像中的边界线长度来测量三维实体中边界的表面积。在其他情形下，如前所述，标记选择或特征“与”逻辑可以用来选择与一个区域邻接（进而接触其卡斯特轮廓）的所有特征。

8.15　欧氏距离图

形态学运算中会出现方向性偏差，因为这些运算针对的是方形网格中的像素，而由各邻域规则和方向导致的距离变化，可通过执行采用不同技术的一种等效运算（腐蚀、膨胀、开运算、闭运算等）来克服。这种等效运算充分利用了由原始二值图像生成的灰度图像，灰度图像中特征内的每个像素被赋予一个值，该值是到最近背景像素的直线距离。这称为欧氏距离图或 EDM。

本章和前几章中给出的多数图像处理函数，要么对连续色调彩色图像和灰度图像操作，产生二值表示，要么对二值图像操作，产生其他二值图像。欧氏距离图是对二值图像操作来产生一种灰度图像的工具。其定义十分简单：前景中的每个点被赋予一个亮度值，该亮度值等于该点到背景中最近点的直线（“欧氏”）距离。相对于包含有限像素的数字化图像，在模拟或连续图像中这很明确。在多数像素图像中，它是特征中的每个像素到背景中最近像素的距离。

为构建 DEM，搜索所有背景像素来找到与特征中每个像素最近的一个像素，并计算毕达哥拉斯意义上的距离，是非常低效且非常耗时的过程。一些研究人员已经实现了一种不同类型的距离图，其中只在几个方向上测量距离。对于方形像素格网，它要么限制在 90° 方向，要么还包括 45° 方向（Rosenfeld and Kak，1982）。这种测量约定等同于在考虑像素是否相接时，是决定使用 4 邻域还是使用 8 邻域。在任何一种情形下，每个像素到其 4 邻域或 8 邻域的距离均当做 1，而不管方向。因此，如图 8.52 所

示，当与正确的毕达哥拉斯距离相比较时，来自一个点的距离图会产生方形或菱形的人为假象，且它是失真的。因为在这些情形下移动有限，因此有时称这些测量惯例为城市街区模型（在四个方向上连通）或棋盘模型（在八个方向上连通）。在各向同性方面，这些版本的距离图与经典的腐蚀和膨胀相比没有任何优势。

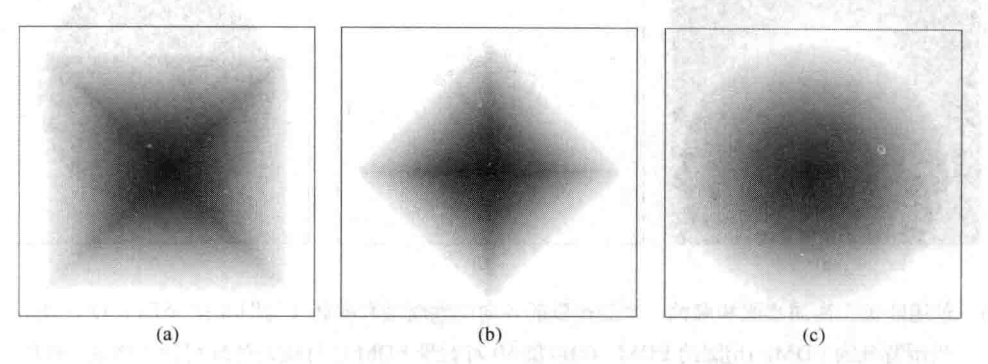


图 8.52 以不同方法测量的到中心像素距离的灰度图像：(a)城市街区或 4 邻域路径；(b)棋盘或 8 邻域路径；(c)毕达哥拉斯距离

构造这样一个距离图的一种简单的迭代技术可以编程如下。

（1）对背景中的每个像素赋亮度值 0。

（2）设置一个变量 N 为 0。

（3）对于与亮度值为 N 的一个像素相接（如上所述，在 4 邻域或 8 邻域的意义上）每个前景像素，赋亮度值 $N+1$。

（4）增大 N 并重复步骤 3，直到图像中所有像素均被赋值。

该迭代（经典的腐蚀）所需时间取决于特征的大小（到背景的最大距离）。一种更有效的方法只需遍历图像两次，就可得到相同的结果（Danielsson，1980）。这种技术使用相同的比较，但在图像中传播值的速度更快。

（1）对背景中的每个像素赋亮度值 0，并对特征中的每个像素赋一个大的正值（大于最大的特征宽度）。

（2）从左到右、从上到下，对特征内的每个像素赋一个亮度值，该亮度值要大于其任何邻域中的最小值。

（3）重复步骤 2，但处理步骤是从右到左和从下到上。

一种进一步的改进可为像素间的毕达哥拉斯距离提供更好的近似（Russ & Russ，2007）。共边的相邻像素间的距离为 1，对角相邻像素间的距离为 $\sqrt{2}=1.414$。对像素使用浮点值而非整数，可为对角相接像素或角点相接像素提供更精确的距离值，并产生改进很大的、更加各向同性的 EDM，即使为显示目的需将最终结果舍入为整数值写回图像内存。

还有另一种算法使用实数值来产生欧氏距离图（Rosenfeld & Pfaltz，1968）。在通过图像期间，对特征内的每个像素，分别累加其到最近背景点的 X 距离和 Y 距离（整数值），然后将实际的毕达哥拉斯距离计算为平方和的平方根。当然，为显示的目的，仍需要将实数结果转换为整数表示。通常，EDM 数值的质量越好，则使用 EDM 进行腐蚀、膨胀和分水岭分割所得的结果就越好，如下文所述。在许多情形下，分水岭分割产生的结果之所以不好，是因为使用了有限的精度和整数算术，导致了 DEM 的质量较差，这是当年计算机性能比目前要差时暂用的方法。当前针对 EDM 开发的许多方法，目的通常是为了提升处理速度（Fabbri et al., 2008）。

对 EDM 进行阈值处理，选取比背景远某个距离的所有像素，可实现腐蚀。膨胀使用背景的 EDM 来选取比边缘远某个距离的像素。使用 EDM 的腐蚀和膨胀方法和逐像素方法的比较（见图 8.38 至图 8.42）

表明，EDM 方法更具各向同性（见图 8.53）。此外，距离图构建的速度更快，且阈值处理不需要迭代，因此该方法的执行时间不会因特征尺寸的增大而增加（如经典的迭代形态学方法一样）。

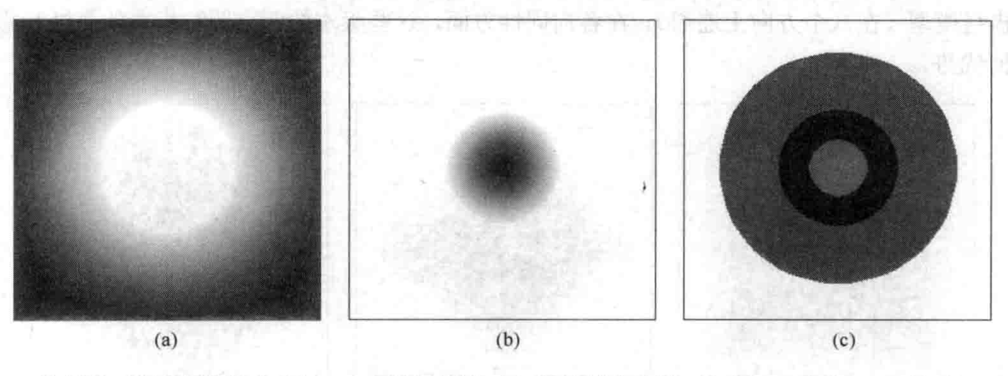


图 8.53 使用欧氏距离图膨胀和腐蚀一个圆得到的各向同性腐蚀和膨胀（与图 8.38 至图 8.42 比较）：(a)圆周围背景的 EDM；(b)圆的 EDM；(c)以值 50 对背景 EDM 进行阈值处理得到的膨胀（蓝色）和以值 25 对圆的 EDM 进行阈值处理得到的腐蚀（红色）

当更多不规则形状受到腐蚀和膨胀时，迭代方法和对 EDM 阈值处理之间的差异也很明显，因为 EDM 方法避免了传统形态学工具中出现的 90° 或 45° 边界。图 8.54 显示了对图 8.49 中的图像应用闭运算和开运算的相同例子。用于闭运算和开运算的距离均为 5.5 像素，且最终的轮廓以更高的保真度跟踪了结构的边缘。注意，使用 EDM 有可能将距离指定为实数而非限制为整数，而且与迭代次数不同的是，EDM 值是实际距离。下面使用 EDM 的这个特性来完成测量。

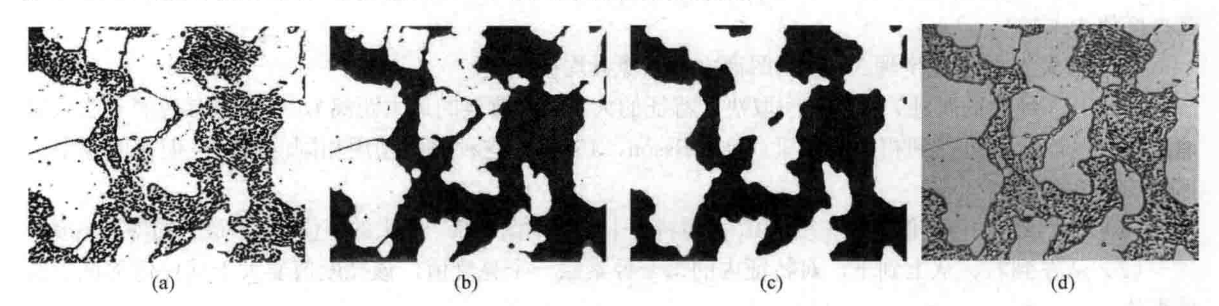


图 8.54 使用 EDM 的闭运算与开运算：(a)对原始图像进行阈值处理后的结果［与图 8.49(b)相同］；(b)距离为 5.5 像素的闭运算结果；(c)距离为 5.5 像素的开运算结果；(d)叠加到原始图像上的轮廓

8.16 分水岭分割

由于物体接触并因此不能单独地识别、计数或测量，因此测量图像时会遇到困难。检查正传送的体积图像或较厚的图像切片，由于可能出现实际的特征重叠，或表面上的颗粒因聚集而彼此接触，因此可能会出现这种情形。分开图像中相互接触的凸状特征的首选方法称为分水岭分割（Beucher & Lantejoul，1979；Lantejoul & Beucher，1981；Sun & Luo，2009）。

完成这一分离的经典方法（Jernot，1982）是一种迭代方法。它是图 8.43 所示例子的逻辑扩展。图像被反复腐蚀，且在每一步中，将那些从前一步消失的单独特征指定为最终的腐蚀点（UEP），并连同迭代次数一起保存为图像。保存这些内容是必要的，因为特征通常有不同的大小，且在相同次数的迭代中不会全部消失。继续该过程直到图像被清除。

然后，从最后的 UEP 图像开始，使用经典的膨胀来膨胀图像，但要增加一个逻辑约束条件，即如

果一个新像素导致先前分离的特征间出现了连接，或如果在原始图像中它不为“开”（非合并扩张），则将其转换为“开”。在膨胀的每个阶段，使用逻辑“或”运算，将对应于等同腐蚀的 UEP 图像添加到图像中。这个过程会使得特征生长回其原来的边界，但相接特征间出现的分隔线除外。

刚才介绍的方法有几个缺点：迭代过程缓慢，需要多次存取图像中的每个像素；腐蚀和膨胀过程中会出现各向异性；存储所有中间图像需要很大的空间。使用 EDM 可以更有效地得到相同的结果。“分水岭”一词源于 EDM 的实现。假设 EDM 中特征内的每个像素的亮度值对应于一个物理高程，那么特征将作为山峰出现。图 8.55 显示了一个这样的圆形特征。

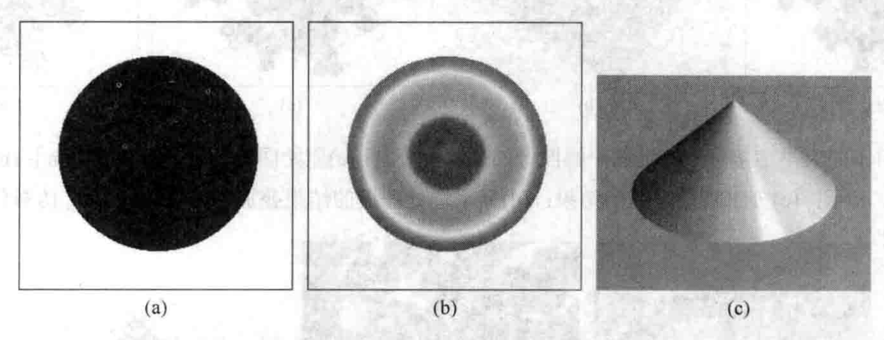


图 8.55　将欧氏距离图解释为像素高度：(a)圆形特征的二值图像；(b)欧氏距离图，已对像素添加颜色来显示到边界的距离；(c)显示像素高度的渲染图

如果两个特征相接或稍有重叠，那么 EDM 会显示出两个山峰，如图 8.56 所示。山坡的斜率是恒定的，因此较大的特征具有较高的尖峰。最终的腐蚀点是山的尖峰和特征相接处（即山两侧的相交处）。这些山之间的鞍部就是由分割方法选为边界的线条。它们是从两座山上流下来的水的汇聚位置，因此称为分水岭线。根据山的相对高度（特征大小）来放置这些线，可最好地估计特征间的分隔线，它根据属于每个山顶的区域来划分。分水岭处理的一个等效描述是，用逐渐充满水的盆地取代山，并用它们之间脊线取代分隔线。

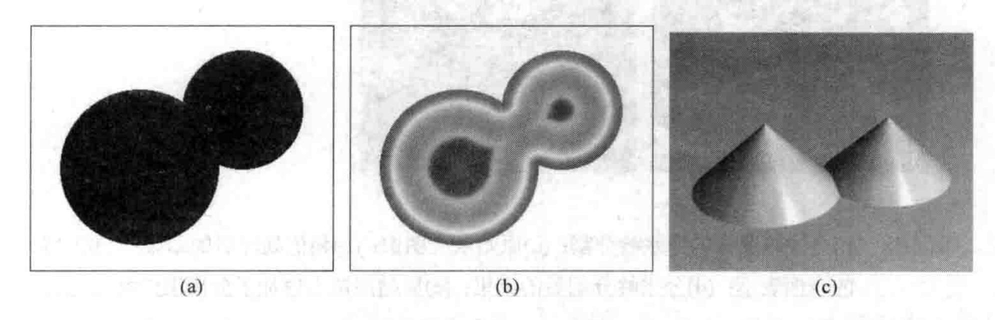


图 8.56　特征相接的 EDM：(a)两个相接圆形特征的二值图像；(b)欧氏距离图，已对像素添加颜色来显示到边界的距离；(c)显示像素高度的渲染图。注意两个锥体间的边界

使用 EDM 方法实现分割过程，在速度和存储效率方面，都非常快（Russ & Russ，1988b）。所需要的距离图图像可以不用迭代的方法构造。最终的腐蚀点定位为局部最大的一种特殊情况（见下面关于 UEP 的深入讨论），且每个亮度值直接对应于因腐蚀而消失的迭代次数。因为距离图提供了一个约束条件，因此对这些特征进行膨胀的速度很快。从最大值开始并“走下山”，涵盖了所有的亮度级。在每一级，只需要考虑距离图中具有当前亮度级的那些像素。将那些不与其他特征相结合的特征，添加到图像中。继续这个过程直到恢复了特征中的所有像素，但沿分隔线出现的那些像素除外。

图 8.57 显示了这种方法的一个例子，这种方法应用到了具有两个相接圆形特征的图像。由于圆形特征的大小不同，因此上面图 8.34 所示的方法不起作用，但分水岭分割分开了这两个特征，并可对它

们进行测量。分水岭分离后，得到圆环的大小分布如图 8.43(d)显示，请将它与迭代腐蚀的结果进行比较。对于一幅真实微粒的图像，如图 8.58 所示，该方法的约束条件是，特征要足够凸，即 EDM 不会在每个特征内产生多个尖峰。

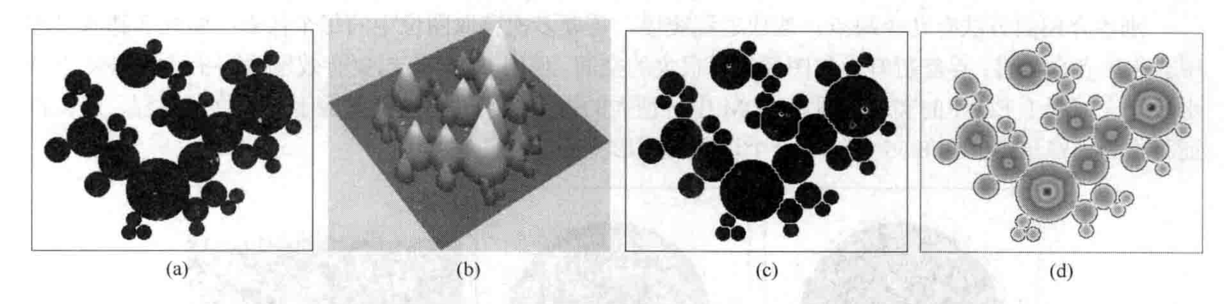


图 8.57　具有不同大小相接圆形特征的一幅图像的分水岭分割：(a)原始图像［与图 8.43(a)相同］；(b)EDM 的“表面”表示；(c)分水岭处理后的结果；(d)分水岭处理后的结果叠加到了 EDM 上，伪彩色表示距离值

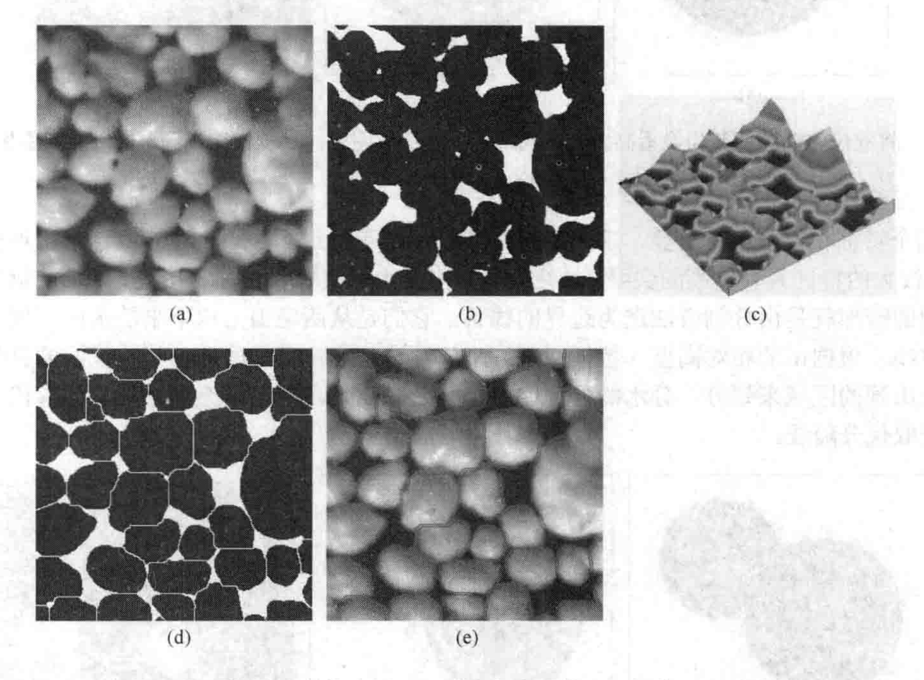


图 8.58　相接沙粒图像的分水岭分割：(a)原始灰度图像；(b)阈值处理后的结果；(c)EDM 的表面表示；(d)分水岭分割后的结果；(e)原始图像上叠加了分离出的特征轮廓

分水岭分割被用做前面几个例子的一部分。当然，这种方法并不完美。分水岭分割不能处理凹状和不规则颗粒（更具体地说，它可以处理非凸颗粒，前提是该颗粒在 EDM 中只有一个最大值），也不能分离那些重叠太大以至于它们之间的 EDM 中无最小值的颗粒。取决于原始距离图的质量，分水岭分割可将固定宽度的线条分成许多小线段，因为沿线边缘的混叠产生了明显的极小。多数情形下，纠正此类缺陷所付出的努力，要远少于人工分离原始特征的努力。

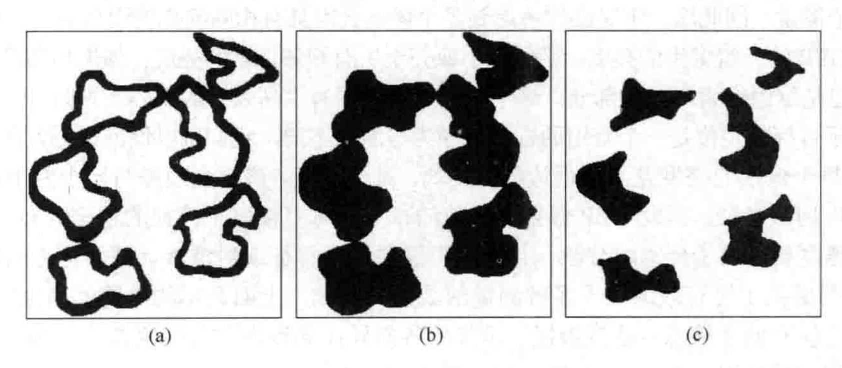
使用其他布尔运算将分水岭分割与形态学过程结合，可以克服这种方法中的许多固有限制。下面的两个例子由很多步骤组成，乍看之下可能过于复杂，但各步骤却很简单且速度非常快，整个运算也是是快速且有效的。

特征内出现的孔洞迷惑了分水岭算法，并将特征分解成了许多碎片。因此，需要在应用分水岭算法之前填充孔洞。但二值图像中特征之间和特征内部还可能存在缝隙。使用常规的填充方法就可填充

这些孔洞，因为与图像边缘不相接的任何背景区域都不考虑为孔洞。如果能找到孔洞大小和形状的差别，进而只填充特征内的孔洞而不填充特征间的孔洞，就可克服这一困难（Russ，1995f）。在图 8.59 所示的例子中，特征（细胞内的细胞器）内的孔洞要比相接细胞间的缝隙圆很多。通过测量隔离出它们，并将它们与原始图像进行“或”运算，就可使用分水岭分割来分开这些细胞。

图 8.59 相接细胞的分离：(a)原始灰度图像；(b)阈值处理后的图像；(c)细胞内孔洞产生的不正确分水岭分割；(d)反转图(b)以显示细胞内和细胞间的孔洞；(e)按形状圆度选取的细胞内的孔洞；(f)使用布尔“或”运算组合图(b)和图(e)；(g)图(f)的成功分水岭分割；(h)分离后的细胞轮廓叠加到原始图像上

图 8.60 显示了分水岭方法的另一个问题。该图像是围绕神经纤维的相接鞘壳的截面。因为孔洞的形状不规则，所以填充孔洞并不能让分水岭方法有效地分开它们。分水岭方法不仅分开了特征，还将



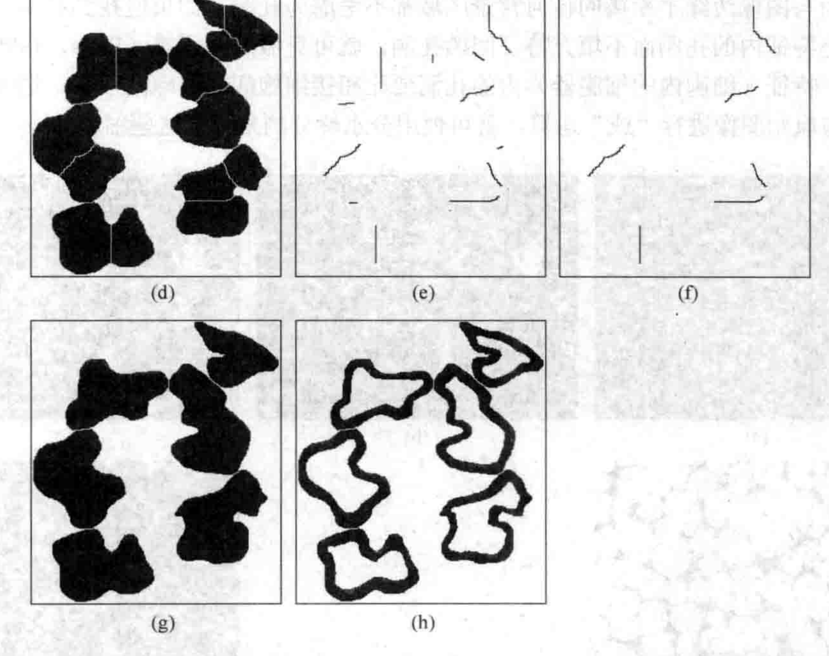


图 8.60　相接不规则特征的分离：(a)相接神经鞘壳的二值图像；(b)填充孔洞后的结果；(c)特征的内部［图(a)“异或”图(b)］；(d)图(b)的分水岭分割；(e)分水岭线［图(b)“异或”图(d)］；(f)图(e)中穿过特征内部的线条［由图(b)中的标记选取］；(g)图(f)“或”图(d)；(h)图(g)“与”图(a)，生成了完整且分离的神经鞘壳

这些特征切成了小片。使用原始鞘壳内部的图像作为一组标记来选取横切特征的分水岭线后，可使用“或”运算使这些线条出现在图像中，进而完成所需的分离。注意，该过程中使用了基于像素和基于特征的两种布尔运算。只有理解了各个步骤，才能理解这些操作。

8.17　最终的腐蚀点

在上述分水岭分割技术中描述的最终腐蚀点（UEP）可直接用做测量工具。点数可给出图像中可分离特征的数量，而每个点的亮度则给出原始特征大小（内切圆的半径）的度量。此外，如果研究聚类或梯度，那么每个点的位置可作为特征的位置。图 8.61 给出了一个例子，并将该测量结果与前面图 8.45 中的连续腐蚀方法所生成的结果进行了比较。

在连续而非基于像素的一幅图像中，UEP 的正式定义是 EDM 图像中的局部最大亮度。由于该图像已被划分为有限个像素，因此这一定义必须考虑到多个像素会因具有相同的亮度而形成“高原”的可能性。换言之，成为 UEP 的一组像素的亮度，必须等于或亮于所有相邻像素的亮度；如果相邻像素的亮度相同，那么它们必须也是这组像素中的一部分。这个位置然后可视为“高原”的质心（见第 10 章中的计算）。UEP 的位置对于目标的定位是一个有用的标记，因为与质心不同，我们可以保证该位置在特征的边界内。

距离图中每个像素的亮度是到最近边界的距离。对于 UEP，该点必须是与至少两个边界等距的点。亮度是特征的内切圆半径，因此 UEP 像素亮度值的直方图可直接度量特征的粒度分布。由于消除了膨胀过程，因此速度要远快于分水岭分割，且由于不需要识别特征或对像素计数，也比测量方法快很多。

欧氏距离图提供了可有效地用于多种测量的数值。例如，上面介绍的由连续腐蚀和膨胀运算来求出分形维数的方法有两个缺点：速度很慢，运算的各向异性导致有方向性偏差。EDM 提供了一种简单的方法来得到相同的信息（Russ，1988），详见第 11 章。

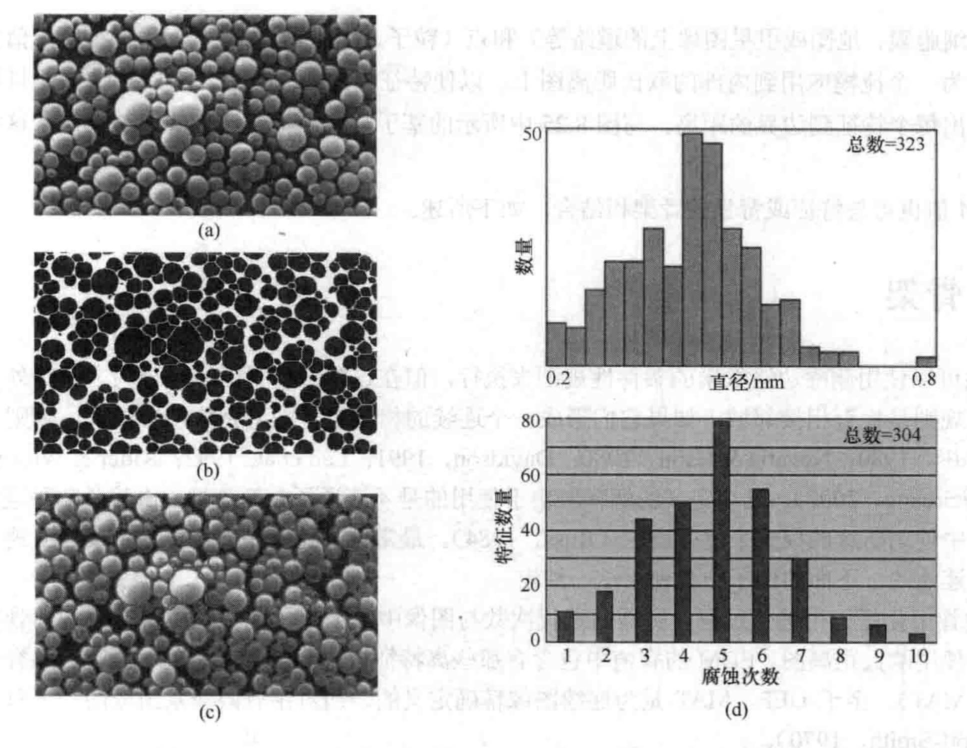


图 8.61　相接颗粒的测量：(a)原始图像（脂肪颗粒的 SEM 图像）；(b)经阈值处理和分水岭处理后的二值图像；(c)使用最终的腐蚀点估算的尺寸和位置（分别显示为红色圆和绿色点）；(d)该过程测量的尺寸分布（顶部，红色）与图 8.45 中对同一幅图像应用迭代灰度腐蚀方法得到的尺寸分布（底部，蓝色）的比较

由于距离图使用每个像素到最近背景点的直线距离来对像素进行编码，因此可用于测量许多点或特征到不规则边界的距离。在图 8.62 所示的例子中，图像给阈值处理后定义了边界线（它可能表示颗

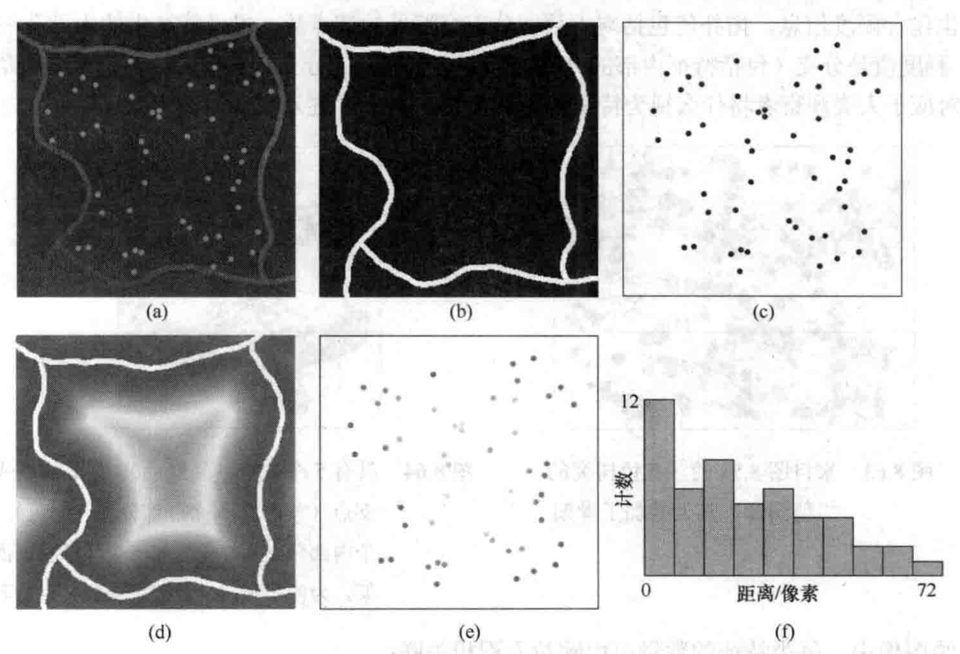


图 8.62　测量到边界的距离：(a)示例图像；(b)经阈值处理后的内部区域；(c)经阈值处理后的特征；(d)内部的欧氏距离图（已着色）；(e)赋给特征的距离值；(f)特征的距离直方图

粒边界、细胞膜、地图或卫星图像上的道路等）和点（粒子、细胞器、结构等）。特征经阈值处理后的图像，作为一个掩模应用到内部的欧氏距离图上，以便特征内的所有像素都有距离值。测量特征的亮度即可得出每个特征到边界的距离。与图 8.25 中所示的基于膨胀和标记选择的方法相比，这种方法更快更准确。

EDM 值也可与特征或背景的骨架相结合，如下所述。

8.18 骨架

腐蚀可以使用删除边缘像素的条件性规则来执行，但在这样做使得区域一分为二时除外。解决这一问题的规则是检查相接邻域；如果它们形成一个连续的相接组，则可删除中心像素，否则就不能删除（Pavlidis，1980；Nevatia & Babu，1980；Davidson，1991；Lan et al., 1992；Ritter & Wilson，1996；Udre & Vizireanu，2007）。这个定义的解释取决于使用的是 4 连通还是 8 连通。不管是哪种连通，都可在命运表中使用所选的模式来进行腐蚀（Russ，1984）。最常见的约定是，特征、骨架是 8 连通的，而背景是 4 连通的，下面的例子中就使用这一约定。

腐蚀骨架化是一个迭代过程，所需的迭代次数与图像中最大尺寸的特征成正比。构建骨架的另一种方法是使用欧氏距离图。EDM 的脊值中包含有那些离特征边界至少两个点的等距点。该脊形成了中轴变换（MAT）。至于 UEP，MAT 是为连续图像精确定义的，而对由有限像素组成的图像只是近似定义的（Mott-Smith，1970）。

多数情形下，MAT 相当接近地对应于顺序腐蚀得到的骨架。由于与任何腐蚀模式相比，MAT 对方向更不敏感，且由于表示一条线的像素有限，因此在某些情形下它可能会稍有移位。MAT 的使用方式与骨架相同，许多情形下会使用 MAT 程序，因为它有更好的各向同性和更快的速度，但结果通常仍称为骨架。

图 8.63 显示了几个特征及它们的骨架（8 连通）。骨架是特征识别的一个强大的形状因子，即包含拓扑信息，也包含测度信息。拓扑值包括端点数、分支交汇处的节点数，以及特征中的内部孔洞数（骨架中的环）。测度值是分支（包括特征内部的分支和具有自由端点的分支）的平均长度和分支的角度。这些参数完全对应于人类观察者将什么视为特征的显著特性。图 8.64 显示了所用的术语。

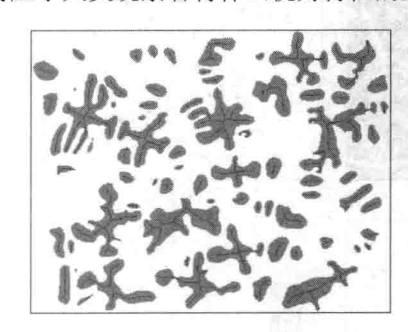


图 8.63　来自图 8.51 的浅灰色树突的二值图像，其上叠加了骨架

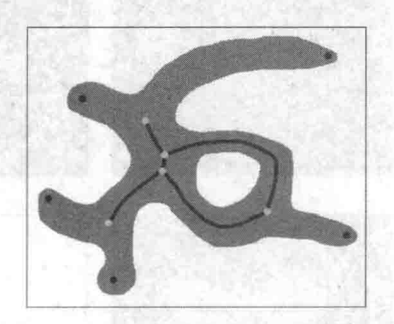


图 8.64　具有 5 个端点（品红色）、5 个节点或分支点（黄色）、5 个外部分支（绿色）、5 个内部分支（红色）和一个环的特征的骨架。为便于观察，骨架已膨胀并进行着色

在二维图像中，每类特征的数量可由欧拉方程相关联：

$$\text{环数} = \text{分支数} - \text{端点数} - \text{节点数} + 1 \tag{8.2}$$

有些特定情形似乎违背了拓扑学的这个基本规则，因此这里需要详细解释数字化骨架。图 8.65 显示两种这样的情形。环骨架化到只有一个圆环、一个单独分支且没有明显节点的单个圆环。但拓扑规则要求环上的某个位置（即连接线性分支两端的位置）有一个“虚拟的”节点。同样，对称圆环骨架化到单个点，该点的邻域少于两个时将分类为端点。实际上，该点表示的是具有两个端点的一个短分支。针对此类情形，可使用特殊的规则。

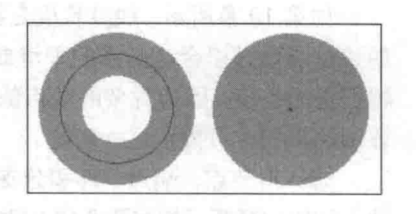
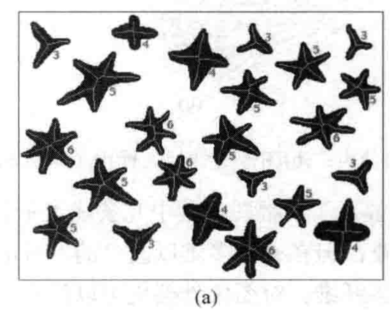
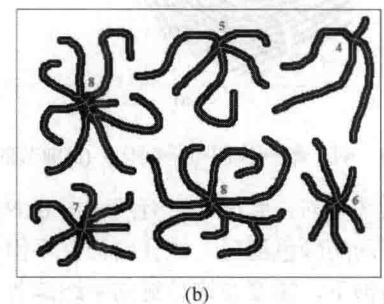


图 8.65　文中所述环和圆的骨架

定位骨架中的节点和端点不过是对相邻像素计数。沿 8 连接骨架分支出现的点恰好有两个相邻像素。端点只有一个相邻像素，节点有不低于两个（3 个或 4 个）的相邻像素。特征的拓扑是一个能从骨架中迅速确定并立刻识别的形状描述子。例如，在图 8.66 中，每个星状物内的点数很容易就可被人们当做定义形状的参数。通过统计只有一个相邻像素的骨架像素的数量，就可使用这种拓扑性质标记它们。当人眼很难识别特征时，这种方法也能起作用（如图中所示）。同样，字母 A、B 和 C 之间的简单区别是环的数量（分别为 1、2 和 0）。按照拓扑性质，它们并不取决于字母的大小、位置或变形（例如使用不同字体造成的）。



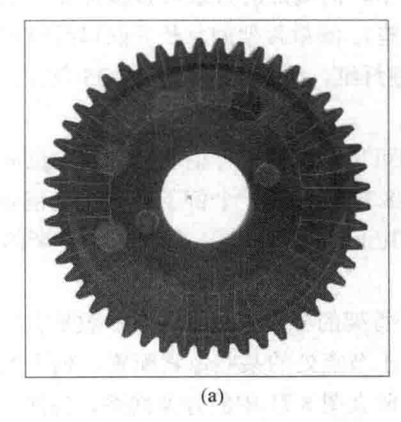
(a)



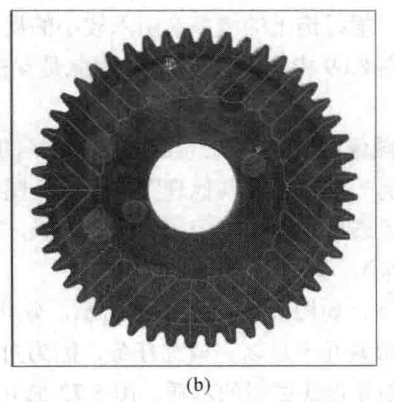
(b)

图 8.66　根据骨架中的端点数量来标记星状特征和分支特征

取决于构造骨架的方式（或 EDM，由它可导出 MAT），它可能会出现方向的各向异性。图 8.67 显示了一幅齿轮图像的骨架。骨架中的端点数（49）立即给出了齿轮中齿的数量。图 8.67(a)所示的骨架是由高精度 EDM 构造的，且有从圆芯引出的径向辐条（在方形像素范围内）。图 8.67(b)中的骨架是由一种广泛使用的免费程序构造的，该程序使用连续腐蚀和整数算术形成了骨架，并生成了具有较强方向性偏差的线条。检查这样一个骨架，通常可测试图像处理软件中所用的底层实现是否良好。



(a)



(b)

图 8.67　齿轮的骨架：(a)由软件 Reindeer Graphics Fovea Pro 软件生成的骨架与原始齿轮图像的叠加；(b)由 ImageJ 软件生成的同样一幅图像的骨架

如第 10 章所示，线度长度是特征尺寸的重要测度。长度测量还可通过统计对角或垂直连接像素对的数量或通过拟合这些点的平滑曲线来实现，此时得到的结果更准确。骨架长度要稍短于由其表示的特征臂的长度，因为骨架的端点像素位于特征臂终点处的内切圆的圆心。对相应像素添加 EDM 值（即该圆的半径）可校正这一问题。

统计出节点、端点、环和分支的数量，就定义了特征的拓扑。这些拓扑事件可简化原始图像，并有助于结构的表征，如在图 8.68 中对指纹图像所做的那样。指纹中的细节用于匹配和识别，并取决于在图案中对断纹、环纹和分支的定位。骨架识别出了它们并使得测量线段的长度和方向（由其可找到纹心）成为可能。



图 8.68　对一幅指纹图像骨架化：(a)细节的分类；(b)图像的骨架，标出了端点和节点

处理交叉纤维图像时，骨架非常有用。图 8.69 显示了一幅图，其中几条线表示相互交叉的纤维。因为这些线条以不同的颜色绘出，统计它们并不困难；但在有许多难以区分的纤维的情形下，计数就会很困难。有些情形下，需要在这种缠结中跟踪各条纤维。对图像骨架化可以解决这一问题（通常使用专门的软件和一些有关纤维性质的先验知识）。然后检查骨架交叉处每个节点周围的区域，分辨出表示纤维延续的分支（Talbot et al., 2000；Beil et al., 2005；Ochoa et al., 2010）。标准通常是，局部方向变化较小，且纤维的宽度、颜色或密度保持一致。当纤维以较小的角度交叉时，骨架通常会显示两个节点，其中一条线段属于两个纤维（如图所示）。直纤维的图像比弯曲纤维的图像更容易理清，如第 10 章所示。

如果所需信息只是纤维的总数和平均长度，那么有可能得到更简单的结果。不管节点或纤维交叉节点的数量是多少，纤维的数量都是端点数量的一半，而端点的数量可直接计算出来（一根纤维的端点有可能位于另一根纤维上的概率会引入较小的误差）。测量骨架的总长并除以纤维的数量可得到纤维的平均长度。在图 8.69 中有 12 个端点，因此是 6 根纤维，而骨架总长是 40.33 英寸，因此平均长度为 6.72 英寸。

对于延伸到图像区域之外的纤维图像，减半的可见端点数仍可正确地求出单位面积中的纤维数，因为每根纤维的另一端点可在其他视野中计数。图 8.70 给出了一个例子。无纺毛毡面料由光纤的密度表征。该图显示了造纸过程中使用的纤维素纤维，其上叠加了骨架。在 6 平方毫米的面积上有 193 个端点（96.5 根纤维），其总长为 68.29 毫米。

在其他情形下，如图 8.71 中的图形所示，分开骨架的分支来测量各个参数是有用的，如长度或方向角。删除节点像素并不足以完成此任务，因为对于节点处的某些像素配置，剩下的分支仍可能是连接在一起的。这是 8 连通逻辑的本质。图 8.72 部分放大图 8.71 中的骨架网络，为便于说明，用于拓扑计数和靠近节点的点已被着色。要分开分支，就必须删除这些相邻的点。这种技术尤其适合于如植物根茎这样的分支结构，只要它们可以铺开来产生二维图像。

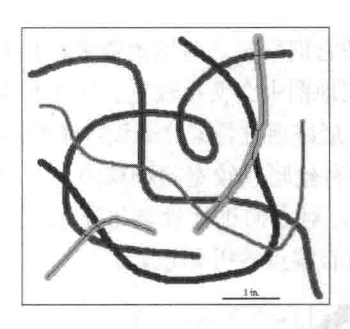


图 8.69　交叉纤维图，其上叠加了骨架线

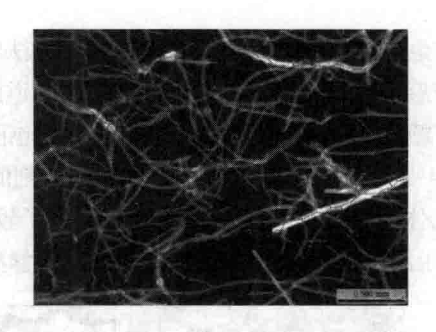
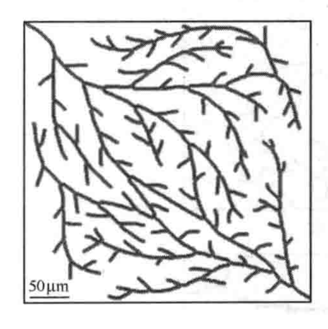


图 8.70　叠加了骨架纤维素纤维图像



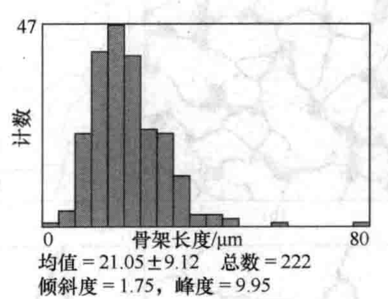


图 8.71　分开网络骨架的分支来测量分支的长度

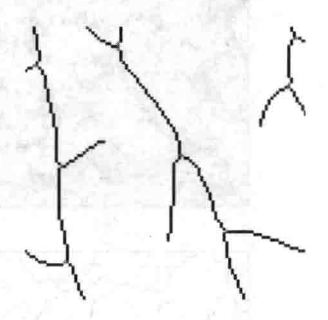


图 8.72　显示有节点的骨架细节。删除红色节点不会断开所有分支；对于某些像素配置，必须删除靠近节点的绿色点，确保不同的分支间无 8 连通残留

正如可确定图像中的特征骨架那样，也可确定背景的骨架。这通常称为特征的“Skiz 骨网”。图 8.73 显示了一个例子。Skiz 骨网由到特征边界等距的点组成，它能有效地将图像划分为几个区域，即每个特征的影响区域（Serra，1982）。我们期望从 Skiz 骨网中删除那些与同一特征两部分边界等距的线条。由于分支有一个端点，因此这种删除很容易实现；Skiz 骨网中的其他线条是连续的，且除了在图像边界处外，没有端点。从骨架（或 Skiz 骨网）修剪分支时，需要从每个端点（有一个相邻像素的点）处开始，并清除相接像素，直到到达一个节点（有两个以上相邻像素的点）。修剪也可用于清理镶嵌，如下所示。

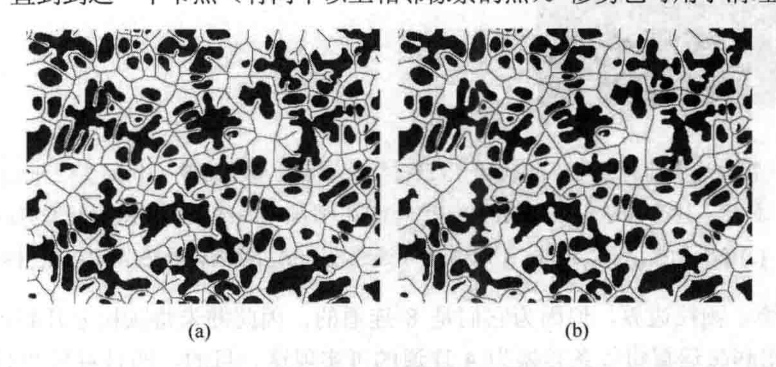


图 8.73　图 8.63 所示图像的 Skiz 骨网：(a)背景的完整骨架；(b)修剪后的骨架

8.19　边界线条和加粗

骨架化的主要用途之一是，使图像中较宽的边界或宽度有变化的边界变细。在金属或细胞的显微图像中，这种现象相当常见，其中化学蚀刻导致了金属的颗粒边界，染色导致了细胞的分界线或细胞

膜。为了产生连续的暗色线条，要采用制备方法和成像来使它们变宽。对地图数字化和处理时，会出现类似的情形，即为让用户看得清楚，边界、道路等在印刷地图中会变得较宽。为了测量细胞的实际大小、不同颗粒类型的邻接关系，或由边界线包围的面积，最好通过骨架化来使线条变细。

图 8.74 显示了一个例子。原始抛光和蚀刻的金属样品有较暗且较宽的颗粒边界，以及对应于碳化物和珠光体的暗斑。对图像阈值处理产生了较宽的线条，对它们进行骨架化处理，可将其宽度降为 1 像素。由于这恰好是连续镶嵌，因此通过修剪删除了带有端点的所有分支。

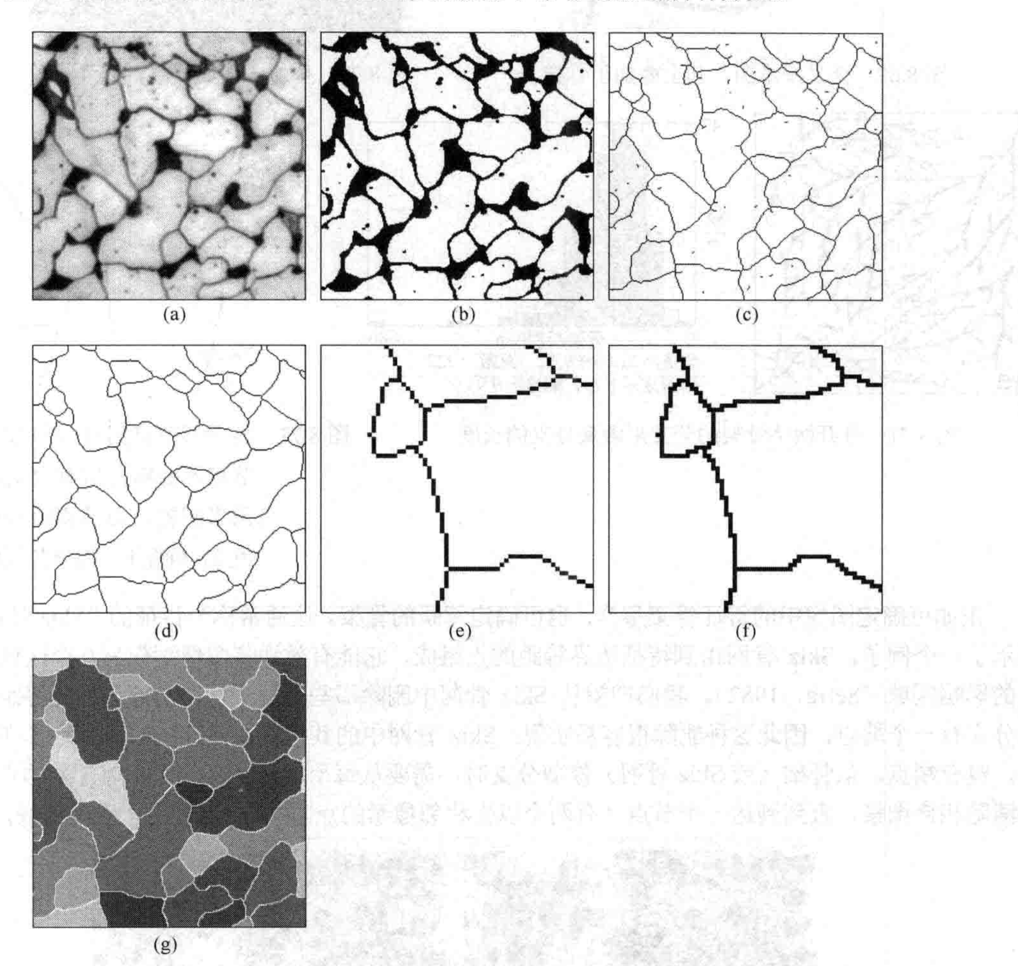
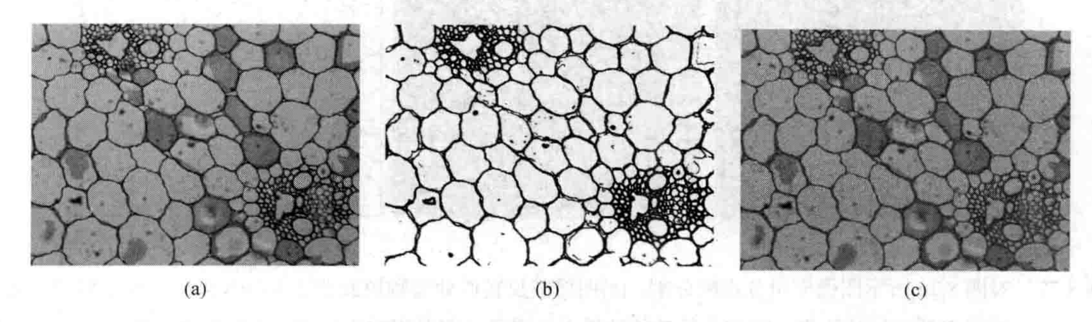


图 8.74 颗粒边界的骨架化：(a)1040 号钢材经蚀刻后的金相图像；(b)阈值处理后的图像，显示了边界和包含碳化铁的暗斑；(c)图(b)骨架化后结果；(d)图(c)经修剪后的结果；(e)放大到显示 8 连通线；(f)转换为 4 连通线；(g)由线细分隔并由指定颜色标识的颗粒

所得线条描绘了颗粒边界，但因为它们是 8 连通的，因此并未将颗粒分开到可对每个颗粒进行测量的程度。使用膨胀运算将线条转换为 4 连通的可实现这一目的，即只对那些对应于 8 连通角相邻图像添加像素（或一开始就可使用 4 连通规则来生成骨架）。所得线条分开了颗粒，进而可以识别和测量。

图 8.75 显示了为测量目的而使用这种方法来分隔图像中的基本结构的方法。原始图像是植物组织细胞的光学显微图像。它可用于测量细胞大小与细胞到两个（开口）气孔距离的变化。通过将细胞壁变为单像素宽的线条，可大大简化这一处理。先调匀原始图像的背景亮度，然后进行阈值处理，

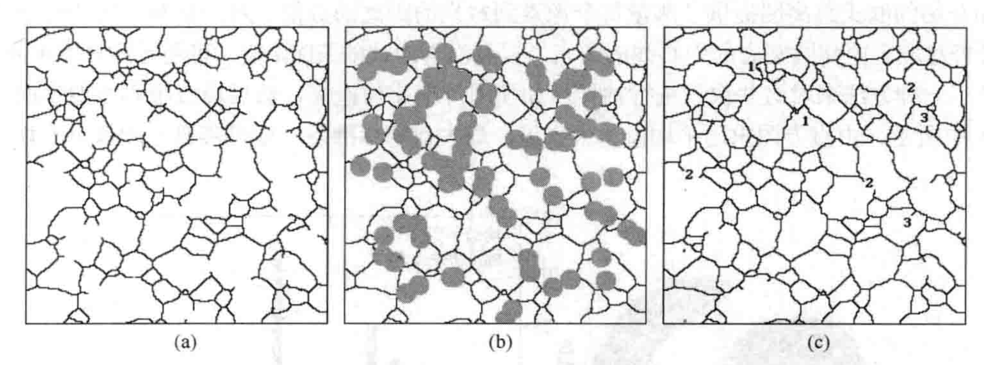
只保留宽度变化的边界线。对这些边界线进行骨架化处理，就可生成能描述基本细胞布局的单像素宽线条网络。



(a) (b) (c)

图 8.75　植物组织细胞的光学显微图像：(a)原始图像；(b)阈值处理后的图像；(c)原始图像上叠加的骨架

遗憾的是，有些情形下简单的阈值处理和骨架化并不能产生完整的颗粒边界或细胞镶嵌。有些边界因与截面的角度问题可能不能成功地染色，或者有些金属可能无法蚀刻，原因是跨越该边界的结晶错位很小，或杂质的浓度较低。因此结果会是缺少了某些线条的镶嵌，而这会使得后续的分析出现偏差。图 8.76 针对这种情形给出了一种方法。对不完整网络进行骨架化，可识别端点（具有单个相邻像素的点）。这些点应成对地出现，因此希望以某个任意选取的距离来膨胀每个点，可修复网络中的一半间隙。



(a) (b) (c)

图 8.76　完成颗粒边界镶嵌的膨胀方法：(a)不完整的网络；(b)以任意半径膨胀每个端点，结果显示为叠加到原始图像上的圆；(c)对网络重新骨架化，显示了某些典型的错误，如删除了较小的颗粒（1），较大的间隙仍未连接（2），悬空的单线（3）

产生的膨胀圆与原始网络相“或”，并对结果再次骨架化。无论膨胀在何处使得圆相接，结果都是连接相应端点的一条线段。但是，这种方法并不完善。有些点可能因太远而使得这些圆接触不到它们，而在其他位置，这些圆可能会因与原有的一些线条相接而掩盖细节，导致得到的网络过于简单。选择合适的膨胀半径并不容易，因为间隙的长度并不一致（且并非所有细胞均有这类间隙）。此外，不匹配的端点、污垢点和细胞内的微粒，同样会造成问题。

还存在其他方法。一种密集计算的方法会找到所有的端点，并使用松弛法来将它们配对，以便保持线条方向，不让线条交叉，并首先匹配较近的点。出现失配端点或噪声时，这种方法也会遇到和膨胀一样的问题，但它能较好地处理不同长度的间隙。第三种方法是基于欧氏距离图（EDM）的分水岭分割，它或许是最有效且精确适中的方法。如图 8.77 所示，它正确地画出多数缺失的线条，但错误地分割了凹状颗粒或细胞（所幸的是，真实细胞或颗粒结构很少是凹状的）。在使用每种方法时，必须要考虑到原始图像的二义性。

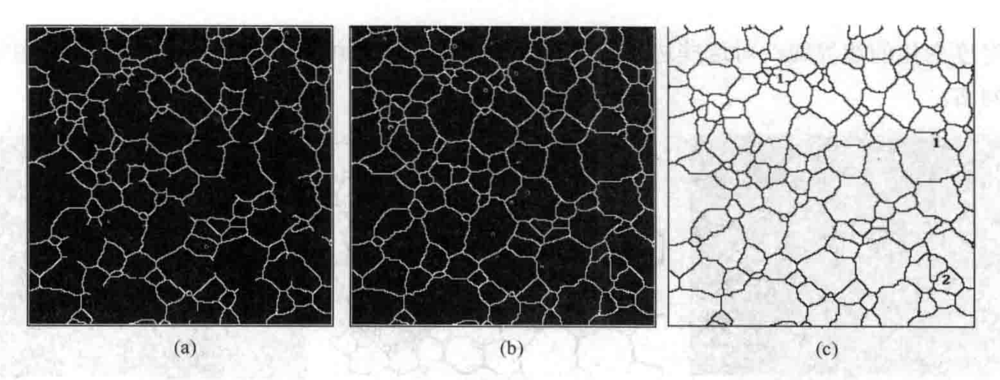


图 8.77　对图 8.76 所示图像应用分水岭分割：(a)图像被反转以处理颗粒而非边界；(b)画出的分水岭线，它连接了多数断开的边界；(c)重新反转的结果中出现了一些典型的错误，如未连接的间隙（1）和不规则形状颗粒的分割（2）

8.20　组合骨架和 EDM

骨架和 EDM 都是测量和处理图像的重要工具，以各种方式组合骨架与 EDM，可有效地提取各种数值来量化图像数据。下面通过几个例子来说明这些技术。

上面介绍的欧氏距离图提供了测量每个像素到背景的距离的数值。对于具有不规则形状或宽度的特征，沿着对应于内切圆圆心的中心线的像素，以及沿着中线的 EDM 值，可用于测量宽度及其变化。骨架提供了一种方法来对这些像素进行抽样，如使用骨架作为掩模，然后检查中心线像素的直方图，如图 8.78 中所示。由直方图求出平均值、最小值、最大值和标准差，便可容易地测量出不规则特征的宽度。

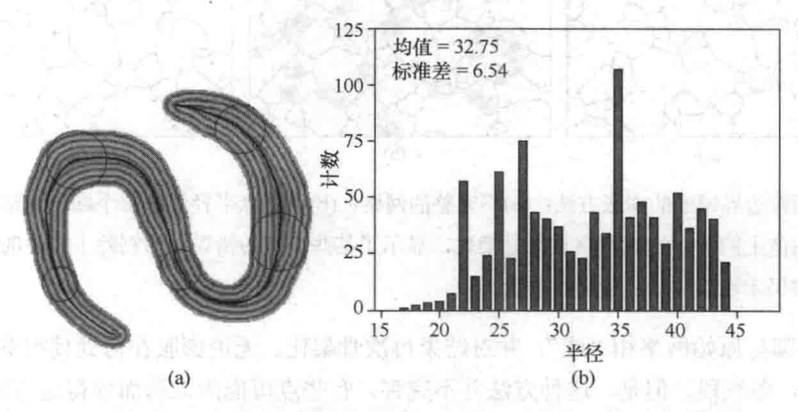


图 8.78　一个不规则的特征：(a)其上叠加了骨架的 EDM，以伪彩色显示，沿中线放置了由 EDM 值定义的一些圆；(b)骨架所选 EDM 值的直方图

骨架是测量此类不规则特征长度的一种基本工具，但骨架长度通常太短。如上所述，校正骨架长度可得到特征的长度，方法是对骨架末端的像素添加 EDM 值。骨架末端像素的 EDM 值给出末端内切圆的半径。将这些值加到骨架长度上，可校正骨架或 MAT 的不足，进而精确地测量不规则形状的长度。

组合背景骨架（Skiz 骨网）与背景 EDM，可求出特征间的最小分隔距离。沿修剪过的 Skiz 骨网的最小 EDM 值，对应于与两个特征相接的最小圆的圆心，两倍的最小 EDM 值对应于分隔距离。

图 8.79 中的例子显示了神经元分支的处理和（从中心孔洞延伸出来的）裂纹的处理。对中央细胞体或孔洞进行阈值处理、反转图像并生成 EDM，就产生了点到细胞体距离的测度。通过删除节点，可

将处理后的骨架或裂纹分到各自的分支中，使用原始骨架端点作为标记，可选取作为末端分支的线段。当这些线段作为掩模应用到 EDM 时，分支中的像素数值就对应于像互到细胞体或孔洞的距离。我们希望使用最小值或平均值来作为有效的距离测度（图中使用了最小值作为到中心的最近距离）。画出每个分支的骨架长度与 EDM 值的关系图后，可看出长度与距离是相关的。

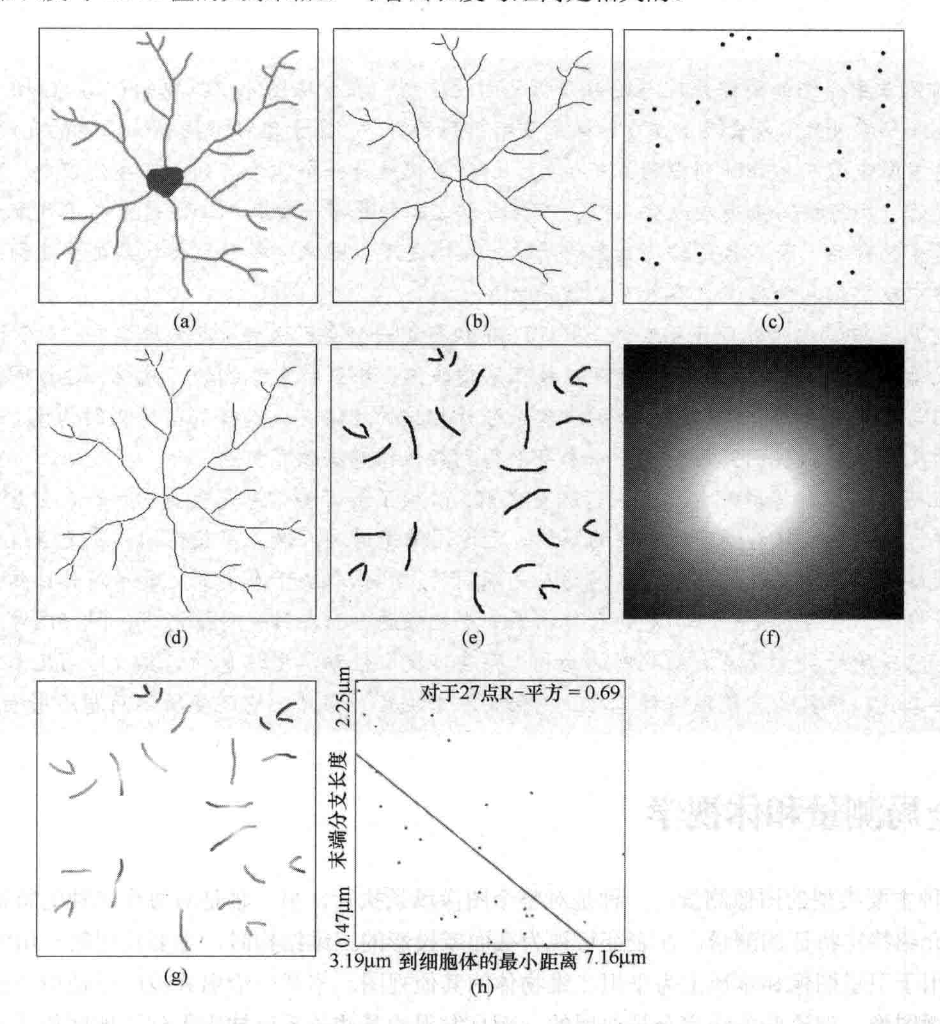


图 8.79　长度与距离的关系：(a)神经细胞图；(b)骨架；(c)末端端点（为便于观察，它已被膨胀）；(d)删除图(b)中的节点后，分开的分支；(e)图(c)中的标记所选取的末端分支，它同样也被膨胀；(f)细胞体外的像素的 EDM；(g)末端分支的骨架线段，已根据其到细胞体的距离着色（这些值按文中给出的方式由 EDM 获得）；(h)距离与长度的关系图

测量出图像中每个特征到 Skiz 骨网上最近点的距离（使用图 8.62 中所示的方法），就可测量图像中的聚类。聚类将在第 10 章中介绍。也可使用 EDM 与骨架的许多其他组合，来求解那些结合了距离和拓扑信息的图像中的测量问题。

第 9 章　全局图像测量

前几章中介绍的图像处理与图像分析之间的区别，在于从图像中提取的信息。如前所述，图像处理如字处理（或食品加工）那样是重新排列的科学。像素值可根据相邻像素的亮度改变，或通过图像变形而移到阵列中的另一位置，但结果仍然是一个像素阵列。在字处理中，剪切和粘贴段落、执行拼写检查或改变字型，不会减少文本的总量。食品加工则是重新排列配料来生产出美味的食品，而不只是转化为配料清单。相比之下，图像分析则试图找到那些能简洁地代表图像中重要信息的描述性参数（通常是数字）。

前几章中考虑的处理步骤在许多情形下是执行这一任务的关键。定义那些要被频繁测量的物体或结构时，需要对图像进行处理以校正采集缺陷，增强特定结构的可见性，通过阈值处理将它们与背景分开，并执行进一步的步骤来分开相接的物体，或选择那些待测的物体。此外，给出的几个程序为我们使用处理方法本身来得到数值信息提供了机会。

在可从图像中提取重要信息进行测量之前，必须了解哪些信息是重要的。如第 2 章中指出的那样，人们往往只看到了所寻找的事物。无论图像是通过人造卫星获得的，还是通过电子显微镜获得的，它所呈现的信息对不同的人是不同的，具体取决于人们到底要了解什么信息。在某种情形下，它可能是一种特定的农作物或一个癌细胞，而在另一种情形下，同一图像给出的可能是建筑物的数量或特定细胞器的面积。存在可提取任意类型信息的工具，后面几章中将介绍这些工具，但具体应使用哪种工具，则仍由人来决定。当然，它还受所问问题的影响。

9.1　全局测量和体视学

有两种主要类型的图像测量：一种是对整个图像场景执行，另一种是对每个单独的特征执行。后面几章将介绍特定特征的测量。在表征可视为截面或投影的三维结构时，主要使用第一组测量，这组测量也适用于卫星图像和本质上为平坦二维物体的其他视图。本章中给出并被广泛使用的分析方法，适用于显微图像，但这些方法完全是通用的，而且发现的某些关系已被应用到了地球科学、医疗成像和天文学领域。

体视学将可在二维图像上执行的测量，与由这些图像表示及采样的三维结构关联起来了。它主要是一种几何和统计科学，其广泛使用的规则和计算非常简单。现代体视学方法的内容可在 Baddeley & Vedel Jensen（2005），Russ & Dehoff（2001），Howard & Reed（1998），Mouton（2002），Kurzydlowski & Ralph（1995），Mandarim-de-Lacerda（2003），Glaser et al.（2007）中找到，经典方法的说明见 Dehoff & Rhines（1968），Underwood（1970），Weibel（1979），Russ（1986），Hilliard & Lawson（2003），Schneider & Weil（2008）。

理解体视学测量的关键是三维结构和穿过它的二维截面之间的关系，就像在显微镜中检查材料和生物标本的情况那样。如图 9.1 所示，截面与三维物体相交生成一个区域，而与表面相交则生成一条线，与线性结构相交则生成一个点。对这些“事件”进行测量和计数，就提供了估计三维结构本身的原始数据。

最简单或许也最常用的体视学处理，是三维体中某些结构所占体积的测量。这可能是细胞中细胞核的体积、金属中的某种相、陶瓷中的孔隙和矿石的含矿量等。体视学家常用“相”来表示他们感兴趣的结构，但它不是化学或热力学意义下的相。如果可以识别图像中的一个结构或相，且如果该图像代表的是整体，那么结构与整个图像的面积比，就是对固体中结构的体积比的测度。事实上，这一关系是体视学中最早知道的关系之一，150 年前就已用于矿产分析中。

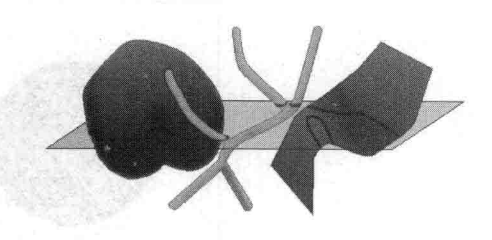


图 9.1　穿过三维结构的一个截面产生了相交。立体（红色）、表面（蓝色）和线性结构绿色）相交分别产生区域、线和点

当然，我们需要对这些假设进行解释和说明。就三维物体的每部分有被检测的同等机会而言，该图像必须具有代表性，以便使截面均匀且随机地穿过三维物体。测量人体中骨骼的体积比时，要求看到头部、躯干、手臂和腿部的机会相同，这是假设的“均匀”部分。随机则意味着不采取任何措施就可使得测量发生偏差，包括或排除图像中的特殊面积。例如，只测量那些具有最少骨骼的图像，会使得测量产生明显的偏差。更微妙的是，在许多情形下，人们倾向于选择那些具有某些感观质量的成像区域（目的在于收集漂亮的图片）。这必定会使得结果出现偏差。正确的体视学采样过程不允许人为地选择或转移视场。

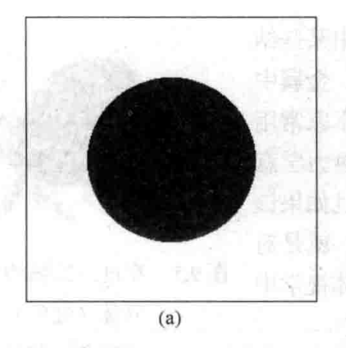
在所有科学领域（特别是显微领域）发表的许多论文中，包含了题为“代表性显微结构”或“典型结构”的图像，这不可能是真实的。要么挑选出的是特定图像，因为这种图像能清晰地显示作者认为重要的某些结构的特征，要么它可给出最佳质量的标本制备和图像对比度，要么某些特性使得它是非典型图像。在多数真实的结构中，不存在真正统计意义上的典型视场。这就是从多个视场收集许多图像并随机且无偏地扩展到整个标本之所以很重要的原因。数据从许多视场中收集，合并这些数据就表示了整个结构。

假设一幅图像是某个结构的均匀且随机的样本，那么该结构面积比的“期望值”等于其体积比。当然，在任何给定的图像中可能并不是这个结果。在有些图像中，感兴趣的相或结构甚至不存在，而在另一些图像中则可能会占据整个视场。通常需要选择一个适当的放大率来使得结构可见且可测，以便检查多个视场并对测量值进行平均。随着包含的测量值越来越多，平均值将接近真实值。

当然，测量结构面积比的一种方法是使用图像的直方图。如果相有唯一的灰度或颜色值，那么直方图中的峰区面积就是所覆盖像素数量的直接测度，进而得出总面积，而不管它占据的是图像中的大区域还是许多小区域。但如前几章中所示，通常要求在对图像进行阈值处理前，进行图像处理，这样才能有选择性地描绘结构，并要求对阈值处理后的二值图像进行编辑。这些步骤也会影响面积测量，因此，在多数情形下必须由最终的二值图像来求面积比。这样做的目的是统计黑白像素。

尽管这个过程很简单，但很难评估测量的准确性。处理沿特征或区域边界出现的像素时会遇到困难，因为阈值处理和随后的形态学处理可能无法包含这些像素，或从总体中排除这些像素。由单个大而紧凑的区域组成的图像，与总面积分布为许多小或不规则特征的图像相比，所具有的边缘像素更少(因此测量误差也较小)，如图 9.2 所示。

有效确定面积比（和体积比）的一种首选方法允许对测量精度进行评估。传统上，这种方法一直采用手动方式执行，但使用计算机也很容易实现。在图像上叠加网点，并统计落在感兴趣结构上的点的比例。这个点比例的期望值也就是体积比。通常，网格中的点数很少，所以人工计数很容易。网格可能位于显微镜目镜的十字线上，或覆盖在荧光屏或照片上，或由计算机生成。这些点应该相隔足够远，以便可单独测量结构（换言之，在正使用的放大图像中，两个点不应落入同一特征中）。如果结构是随机的，那么可以用任何一种网点，规则的方形网格较为方便。如果结构有一些不规则性或周期性，那么网格本身应该是随机的，以防止偏差。



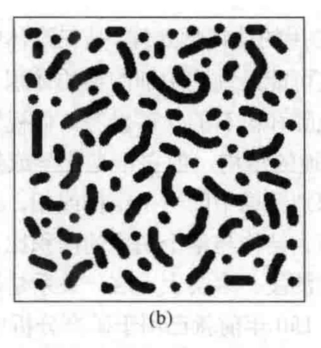


图 9.2　具有相同黑色像素面积比（27%）的两幅图像：(a)一个大而紧凑的区域；(b)许多不规则的小特征。测量精度取决于黑白边界的外围。边缘像素数（邻近白色背景的黑色像素）在图(a)占黑色总面积的 3.4%，而在图(b)中占黑色总面积的 36.1%。注意，大部分观察者都不能非常准确地估计面积比，也无法判断出这两幅图像有相同面积的黑色像素

当网格叠加到（已通过处理得到增强的）原始灰度或彩色图像上后，观察者可以使用独立的知识和判断来确定每个点是否位于感兴趣的区域内。人们非常擅长于这种识别，但并不善于计数，因此标记出这些点并让计算机来计数仍是有用的，如图 9.3 所示。

让这些点成为三维微观结构内的探针，并统计"击中"感兴趣相的点数，就可估计测量的精度，因为对于独立事件，标准差正好是计数的平方根。因此，我们可以快速估计出具有期望测量精度的图像数量（即多个截面上的多个视场）。例如，如果使用 35 点的网格（7×5 阵列，如图所示），且平均有 11 个点位于感兴趣相上，则其对应于 30%左右的体积比。为确定相对精度（如 5%）的体积比（即 30±1.5%），只需将该网格应用到多个视场，直至记录了总共 400 个点（400 的平方根是 20，或 5%）。这应需要 36 个视场（400/11）。这是设计实验过程的正确方法。

使用布尔逻辑运算，也可将点网有效地应用到经阈值处理后的二值图像，如第 8 章所述。如果使用布尔"与"运算来组合网格与二值图像，并统计出剩下的点的数量，则结果刚好是落入感兴趣相上的那些点，如图 9.4 所示。这些关系的常用符号表示是

$$V_V = A_A = P_P \tag{9.1}$$

它表明，材料单位体积内感兴趣相的体积 V_V，等于面积比 A_A 或点比 P_P；更准确地说，单位体积中感兴趣相的体积比可由面积比或占比来测量。

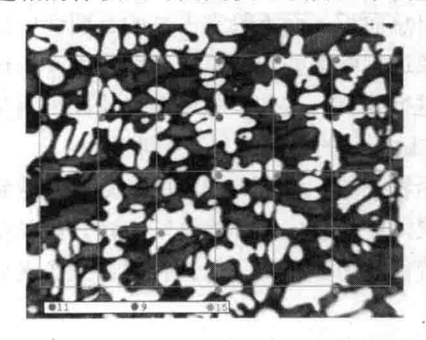


图 9.3　标记的计算机计数：由人将彩色标记放置到靠近每个感兴趣相的网点处，并用计算机对其计数。对于白相（绿点），估算出的体积比是 15/35 = 42.8%

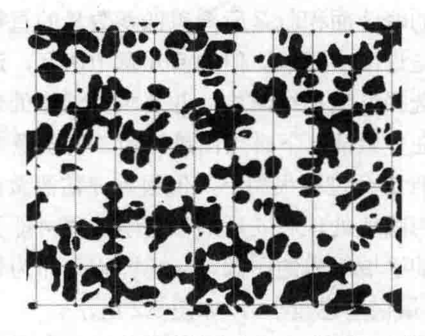


图 9.4　图 9.3 中金属结构的白相二值图像，其上叠加了网格（为便于观察，点已被膨胀并着色）。网点与二值图像的布尔"与"运算只保留那些位于该相上的点（橙色所示），然后用计算机对这些点计数可得 P_P = 33/80 = 41.3%。图中所示点的密度对于最佳精度估计太高，因为存在多个点落于同一特征中的情形。这会影响精度估计，但不会影响测量准确度

体积比是一个无量纲比值，因此无须知道图像的放大率。在特定结构上切一系列截面，然后测量每个截面的面积，也可求出该结构的体积（Gundersen，1986）。如图 9.5 所示，采样空间的体积由截面的大小和间距来定义，由网点形成的“击中”数可绝对测量特征的体积。每个网点对结构进行采样，并表示了空间中的体积等于网格面积的平方乘以截距。当然，在这种情形下必须校准放大率。

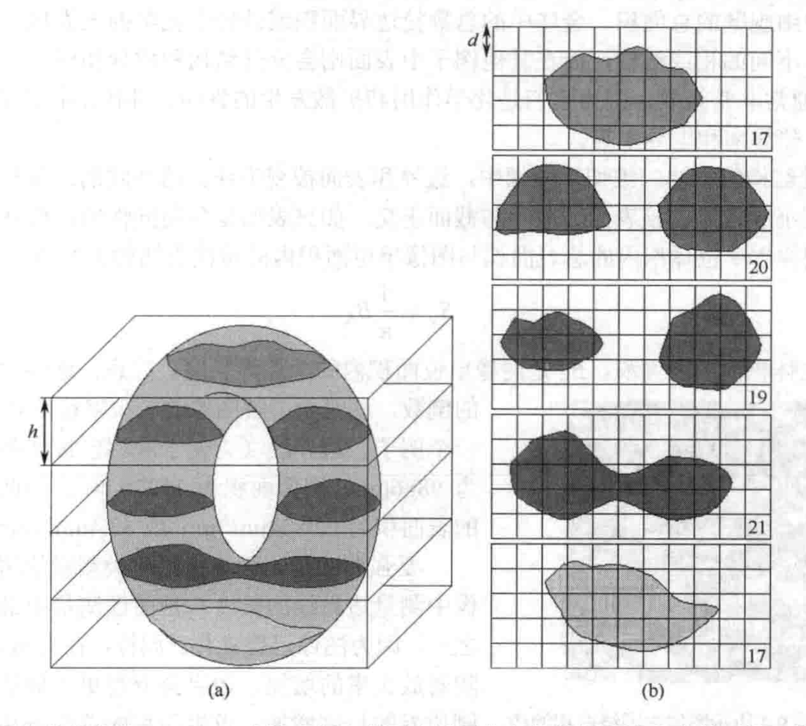


图 9.5 测量物体体积的 Cavalieri 方法。每个截面上的网格把物体划分成几个单元格。统计落在感兴趣结构上的网点，可求出该结构内的单元格的数量，据此可测量体积。在该例中，若截面间距（h）是 5μm，网格间距（d）为 2μm，则每个单元格的体积为 $5×2×2 = 20μm^3$，由击中总数（17 + 20 + 19 + 21 + 17 = 94）可估算出总体积为 $1880μm^3$，94 次击中的估计精度为 $±194μm^3$。测量每个相交面的面积并使用辛普森法则来进行体积分，可得到估计值为 $1814μm^3$

在许多情形下，如图 9.6 所示，需要测量两个体积，即感兴趣结构的体积和一个封闭结构的体积。该例对老鼠肺部截取了一些截面，将每个截面中肺部组织的面积（包括内部空隙）加到一起，可求出净面积。统计像素数量或使用合适的网格，可测量出该面积。肺内空隙面积计算如下：

$$体积 = \frac{\sum 填充的面积 - \sum 净面积}{\sum 填充的面积} \tag{9.2}$$

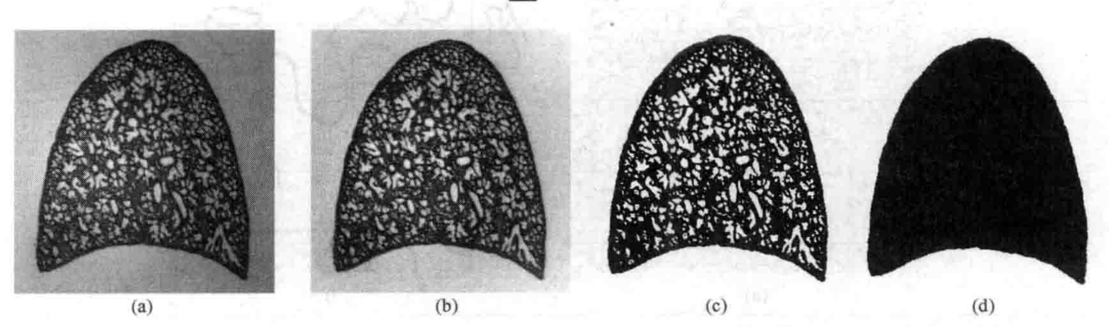


图 9.6 老鼠肺部组织的一个截面：(a)原始图像；(b)对比度调匀后的图像；(c)阈值处理后的图像；(d)填充内部空隙后的图像。如文中所述，图(c)和图(d)的面积用于求肺内的空隙体积

9.2 表面积

体视学测量的另一个结构参数是表面积。表面通常是两个三维物体之间的边界，如细胞核的表面积、植物组织中细胞壁的总面积、金属中的总颗粒边界面积或砂岩中孔隙的表面积。在有些例子中，表面会分开两个不同的相或结构，而在其他例子中表面则会分开结构和成分相同的不同区域（如细胞或颗粒）。表面通常非常重要，因为它们是化学作用和扩散发生的界面，并控制着许多性质，如强度、断裂、光散射、热导率和电导率等。

在穿过三维结构的一幅二维切片图像中，边界和表面被视为线。这些线的总长与三维结构体中出现的表面积的大小成正比。该表面一般不与截面正交。如果表面是各向同性的，或截面关于表面的方向是随机的，则样本单位体积内的总表面积与图像单位面积内的总线长间的关系为

$$S_V = \frac{4}{\pi} B_A \tag{9.3}$$

式中，S_V 是单位体积内的表面积，B_A 是图像单位面积内的总边界长度。注意，这两项的单位均为长度的倒数，这对于了解图像的放大率很重要。图 9.7 显示了一个例子。边界线（为便于观察已被膨胀和染色）的总长为 9886μm，图像面积为 15270μm^2。因此得到单位体积内的表面积为 0.0824μm^2/μm^3 或 82.4mm^2/mm^3。

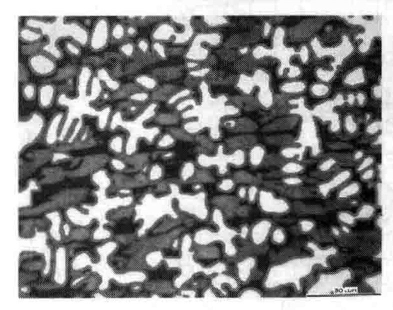


图 9.7　图 9.3 和图 9.4 所示图像中围绕白相的边界线，可使用第 8 章的技术分离与测量，进而根据文中的说明求出表面积（为便于观察，这些线已被膨胀并着色）

要强调的是，很难指定此类测量的精度。在数字化图像中测量边界线的长度，是图像测量中最容易出错的任务之一，因为图像已像素化。同样，在有些结构中可观察到，随着放大率的增加，边界会变得更不规则，同时会使得所测边界的长度增加。求表面积的首选方法是，在图像上放置网格线，并统计网格线与表示相交面的线的交点的数量。单位体积内的表面积（S_V，平方微米/立方微米）与交点数量（P_L，交点数/微米网格线长）之间的关系是

$$S_V = 2 \cdot P_L \tag{9.4}$$

式中，因子 2 是对网格线以不同角度与边界相交的补偿。由于这只是一个计数实验，因此可由交点数的平方根来估计测量精度，前提是这些线（此时是深入到微观结构中的探针）间的距离大到足以使得交点为独立事件。

在图 9.7 所示图像上生成网格线并统计交点数（见图 9.8），可产生与边界线类似的测量结果。网格线的总长为 4646μm。执行布尔“与”运算并统计交点数后，得出 197 个击中（有些是单个像素，而

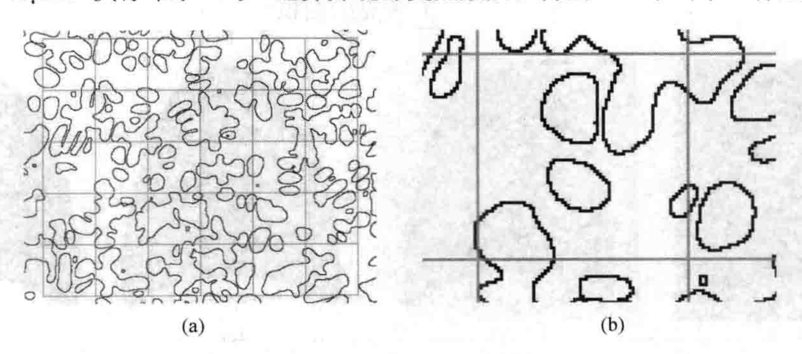


图 9.8　如文中所述，对网格线和图 9.7 中的边界线进行“与”运算，然后统计交点数，也可测量表面积。有些交点会覆盖多个像素，如图(b)中的放大细节所示，但这些交点按单一事件统计

有些是多个像素，具体取决于边界与线之间的角度，但统计的是相交事件而非像素）。由式（9.4）可知，这对应于 0.0848μm²/μm³ 的单位体积内的表面积。基于计数数量，估计出的相对精度是±7.2%（0.0848 ± 0.0061μm²/μm³ 或 84.8 ± 6.1mm²/mm³）。

按照上面统计点数的相同方法，可将网格线放到图像上并统计相交线的数量（手工标记并使用计算机计数，或者使用布尔逻辑运算自动地计数）。如所描述的那样，这种方法的问题是，它依赖于待测表面是各向同性的这一假设，或截面和网格线关于表面的方向是随机的这一假设。在真实结构中，这种假设基本不成立。因此，若结构不是各向同性的，则需要构建不对结构各向同性地采样的网格线（但仍需要上面提及的均匀和随机采样）。现代体视学的许多开发旨在找到一种实用的方法来满足这种各向同性、均匀性和随机性（IUR）需求。为形象化各向同性方向的含义，我们考虑图 9.9 所示的半球。空间中的每个方向由球体上的一个点来表示。这些点应均匀地分布在球面上。

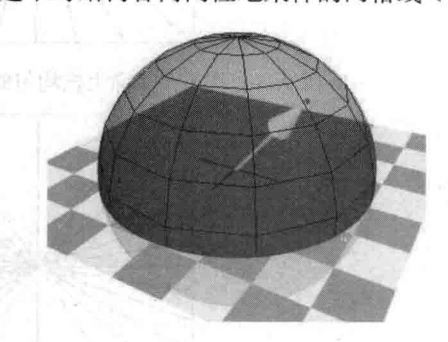


图 9.9 一个半球，它表明三维空间中的方向由球体上的一个点来表示

随机化截面关于结构的方向的一种方法是，先将材料切分为许多小片，然后使用机械手段来随机地对它们定向（例如，在表面上滚动它们或在液体中混合它们），再对它们切片。这样做并非任何时候都实用，且通常而言会造成材料的浪费。

产生各向同性、均匀和随机的线条进而探测三维结构中的表面的首选方法，称为垂直切割（Baddeley et al.，1986）。这种方法需要在总能被识别的结构中选择某个方向。它可能是动物脊椎的轴线，或轧制金属片表面的法线，或岩层的重力方向。然后平行于该方向切割截面，此时要均匀且随机地围绕该方向旋转，如图 9.10 所示。嵌入生物组织并显微切片时（就像从切片机中出来的香肠片那样，这些截面都与同一个方向正交），这并不是切割多数截面的方式。也可像切割一张馅饼那样，使用径向切片法切成垂直截面，与馅饼外侧相比，这会对馅饼的中心过采样。

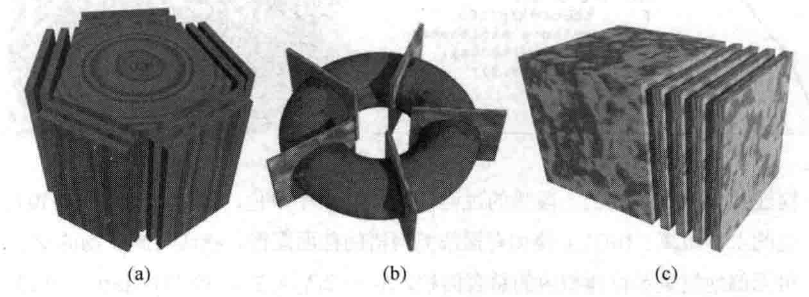


图 9.10 切割垂直截面的方法：(a)和(b)正确，所有截面均平行于所选轴并围绕该轴旋转；(c)错误，截面均与一个共同方向正交

垂直轴方向出现在所有截面上，因此这些截面是有方向偏差的。如果在该平面上使用均匀的方向画出这些截面上的线条，如图 9.11 所示，则它们在半球的北极附近汇聚，而在靠近赤道的方向欠采样。这种偏差可通过画出正弦加权线来补偿，如图 9.12 所示。也就是说，不以均匀的角度间隔画出这些线，而以均匀的角度正弦值间隔画出这些线。这会在赤道附近产生更多的方向并将方向展开到北极附近，使得这些点在球面上均匀地分布。然而，尽管图 9.12 所示的径向线提供了各向同性的方向，但它们不能产生均匀的采样（与角点相比，每个截面的中心会被过采样）。

画出可对区域均匀采样的正弦加权线的最简方法是，生成摆线，如图 9.13 所示。摆线是一个点在滚环圆周上行进的路径。它是正弦加权的，且在垂直截面上画出时，会有完全正确的方向集来在三维体中各向同性地采样。统计图像中由摆线网格与边界线形成的交点的数量，就可得出计算单位体积内的表面积所需的 P_L 值。摆线的每个 1/4 弧长恰好是其高度的两倍。

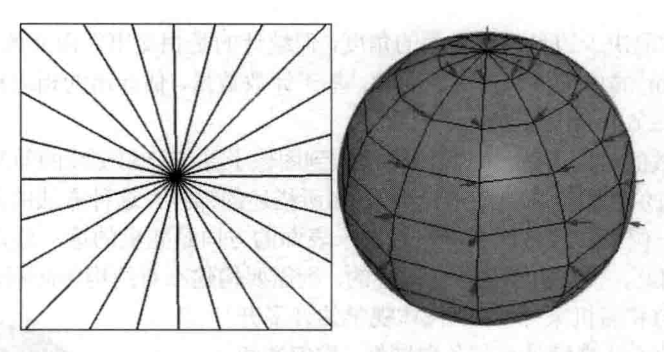


图 9.11　在每个平面上按均匀的角度画出线条，这些线条会在靠近半球的北极处汇聚

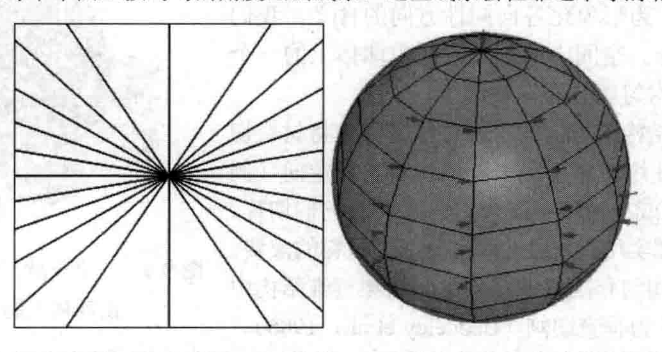


图 9.12　在每个平面上按正弦加权角画出线条，在三维空间中这些线的分布方向是均匀的

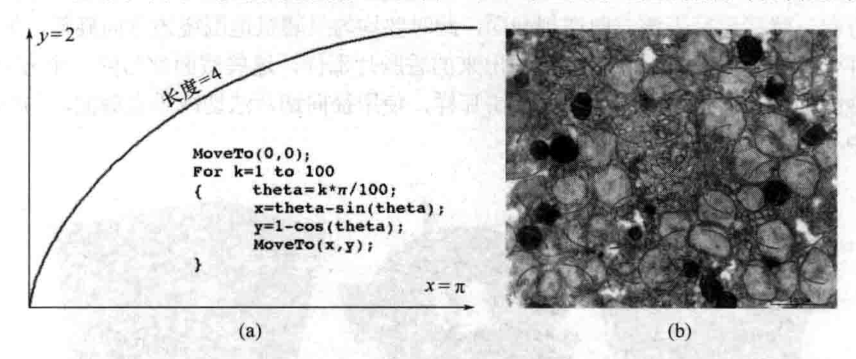


图 9.13　摆线：(a)画出摆线的一段弧的过程，其长为 4 个单位，高为 2 个单位，覆盖了π个单位的水平距离；(b)其上叠加有摆线弧网格的截面图像。摆线与膜状物的交点（红色）可无偏地测量单位体积内的总表面积，$S_V = 2P_L = 2 \times 39/331.5\mu m = 0.235\mu m^2/\mu m^3$

若标本实际上是各向同性的，那么切割垂直截面并用摆线网格来统计交点数量的方法，产生正确结果所需的工作量会比事先已知各向同性时多一些（主要体现在样品制备上）。但是，如果标本有任何各向异性，那么切割平行截面并画出直线网格的方法会产生偏差量不明的错误答案，而垂直截面方法可给出正确答案。

人们已开发出了许多新技术来使用有向截面和线条，包括体视学测量中的随机线和采样点线（Cruz-Orive，2005；Howard & Reed，2005）。这些技术提供了测量任意物体（不必是凸形体）体积和表面积的方法。多数方法的主要问题在于难以实际实现。在使用共焦显微镜对透明矩阵样本成像时，或使用计算机断层扫描（CT）或核磁共振成像技术（MRI）对一系列截面成像时，可以实现各向同性的“虚拟”探针，如三维数据集中的摆线或球面（Gokhale et al.，2004；Kubinova & Janacek，2001）。例如，球体在每个平行平面上显示为一组大小可变且对齐的圆。要特别强调的是，现代体视学是得到无偏采样或任意结构的一种直接方法。

9.3 ASTM 颗粒尺寸

100 多年前，当海军将越来越庞大的火炮引入战舰上时，炮手是使用黄铜弹壳来填装火药并开火的，这种炮弹实质上是现代步枪子弹的放大版。据观察，开火后有时很难从炮膛取出已用过的铜壳，因为在它从后坐边缘拉出时材料已断裂。引爆时，产生的高压会使得金属流动并变薄，因此材料结构方面的某些差异与断裂有关。

根据这一观察，人们开发了一种用于测量显微结构参数的方法，即 ASTM（美国材料和试验协会）“颗粒尺寸”方法，主要用于船用黄铜的质量控制，后来应用到了许多其他的金属和非金属上。实际上，有几种不同的标准技术，它们已用适当的常数进行了调整，以便近似地符合于彼此的数值。但事实上，这些方法测量的是两个完全不同的显微结构特征（二者都不是颗粒的“尺寸”）。

第一种方法统计 100 倍放大率下抛光截面上每平方英寸内的可见颗粒数量。颗粒数 N 通过如下公式与“颗粒尺寸数”相关联：

$$N = 2^{(G-1)} \tag{9.5}$$

式中，G 值的精度不会好于 0.5。

求 G 值的第二种方法是在图像上画出网格。由于形成过程的影响，很多金属是各向异性的，颗粒会在某个方向上拉长。为避免测量的方向偏差，使用了一个圆形网格。在线的单位长度（N_L，单位为毫米）上，线和颗粒边界交点的数量计算如下：

$$G = 6.6457 \cdot \lg N_L - 3.298 \tag{9.6}$$

式中的常数用于转换到公制单位并使得结果与颗粒计数方法基本一致。再次声明，结果被舍入到最接近 0.5；图 9.14 中给出了一个例子。

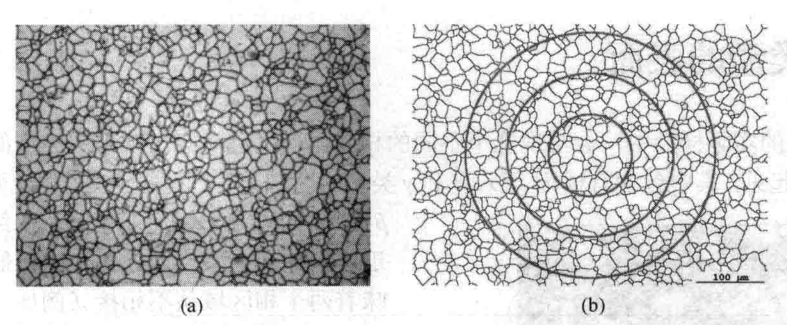


图 9.14　颗粒尺寸测量：(a)低碳钢的光学显微图像；(b)经背景调匀、阈值处理和骨架化后的结果，它描绘了颗粒边界，并叠加了网格。统计最大圆内的颗粒数量（445，与圆相交的每个颗粒按 1/2 个计算）得出了颗粒尺寸数量 $G = 8.8$。用同心圆网格统计截距（2300μm 的网格长度为 147）得出颗粒尺寸数量 $G = 9.1$。这两个值都将颗粒尺寸数量报告为 9

ASTM 标准（E112）中还有一些其他的方法，如统计单位面积内的三相点（三个颗粒交汇的位置）的数量，这些方法几乎都可由手工计数或计算机实现。但这些方法均不能正确地处理由两个颗粒群落（一个较粗而另一个较细）组成的二相显微结构。

截距计数方法［见式（9.6）］实际上是测量颗粒边界的总表面面积。单位长度网格线上的交点数量 N_L 与式（9.4）中的 P_L 值相同，且与单位体积内的表面积成正比。从机械性能来看，颗粒边界面积很有意义，因为颗粒边界滑动和颗粒旋转会导致材料蠕变，而错位运动会导致颗粒变形。颗粒计数［见式（9.5）］和三相点计数方法，实际上测量的是多边形颗粒边缘的单位体积的总长，这个总长是一个主要与扩散相关的参数。下面讨论结构长度的测量。

这些结构特性之一甚至可能近似地与其他结构特性相联系的事实，是由于有限制地对经热处理金属应用这种方法造成的，因为经热处理过的金属会充分再结晶并达到热力学平衡状态。此时，实际的三维颗粒尺寸分布接近一个常数，近似为对数正态分布，即结构会随时间变粗（大颗粒增大而小颗粒消失）但保持为自相似的。因此，颗粒边缘的长度和它们的表面积之间有近似一致的关系。

这种一致的关系并不适用于未经热处理的金属，因此使用单个数值来表征显微结构的说法不正确，即使是对于质量控制目的，它也与所用的测量方法有关。但 100 年的使用认可了该技术，许多从业者并不了解他们的程序在显微结构中实际上测量的是什么，或者说肯定不是字面意义上的颗粒尺寸。

测量三维体中各个颗粒或细胞的“大小”（通常是体积）很困难。有些情形下，这是按照字面意义，将材料分开（例如，沿边界化学溶解一个薄层使颗粒分开）并对每个颗料进行测量或称重来完成的。其他研究人员使用了一种连续切片技术，这种技术可生成结构的三维立体图像，通过该图像即可进行颗粒大小的测量。这两种方法都太实验化而不适于实际应用。使用下面介绍的 Disector 方法，可求出细胞的平均体积。这种方法测量单位体积中的特征数，该数的倒数就是每个特征的平均体积。这种方法在处理大量图像时，严重依赖于基于计算机的图像处理。

对 IUR 剖面应用采样点截距法（见下文中的描述），也测量特征的颗粒尺寸的方差。但是，如果需要实际三维特征的颗粒尺寸，则必须采取更密集的方法。

金属的“颗粒尺寸”只适用于这样一种情形，即特征尺寸的测量与我们所说的普通测量完全不同。例如，在肺部组织切片上画出平行网格线，并测量网格线交点与空气组织边界间的平均距离，可得到一个平均弦长，这已被证明与肺气肿的严重程度相关（Rosenthal & Begum，2005）。弦长看起来似乎代表肺内气道的某些特征大小，但由于它是式（9.4）中 P_L 值的倒数，因此也是对整个肺泡表面积的测量（这涉及气体交换效率）。对砂岩中的孔隙使用弦长分布来描述控制石油流量的孔喉大小，进而部署钻井位置，同样是有问题的。

9.4 多种类型的表面

在多数类型的实际样品中，会出现几个不同的相或结构，进而会出现不同类型的边界表面。在三相材料（为方便起见，其中包含有标记为 α、β 和 γ 类型的区域）中，可能有 6 个相交面（α-α、α-β、α-γ、β-β、β-γ 和 γ-γ），而在更复杂的结构中这一数量会迅速增加。某些类型的相交面可能并不存在，这意味着两个相区域从不相接（例如，不同细胞的细胞核互不相接，考虑两个孔之间的接触面毫无意义，等等）。测量不同类型相交面的数量，对表征整个三维结构很有用。

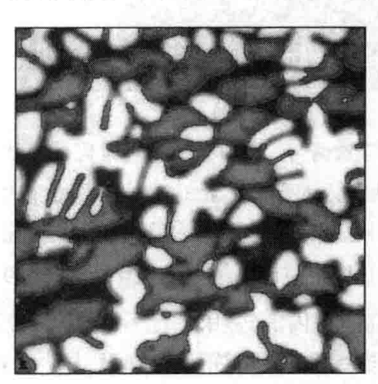


图 9.15　图 9.3 所示三相铝锌合金，突出了区域边界线并用颜色编码，如在第 8 章中的图 8.51 所示

图 9.15 显示了一个三相金属结构中的相边界。注意，白色的枝状区域互不相接。第 8 章中图 8.50 所示的膨胀区域轮廓的布尔逻辑运算，可用于描绘每种类型的相交面。隔离出边界图像后，测量线长或统计与网格的交点数，就可为每种类型的相交面求出单位体积中的表面积。

图 9.16 显示了一个简单的、理想化的二相显微结构。通过对每个相进行阈值处理，使用形态学运算（如闭运算）清除公共区域之间的边界，并用布尔逻辑运算来合并两幅得到的图像，可隔离出三类不同边界中的每一类边界（为便于观察，图中它们以不同颜色显示）。假设这是一个垂直截面，使用式（9.4），那么可用图中所示摆线网格来估计单位体积中的表面积。

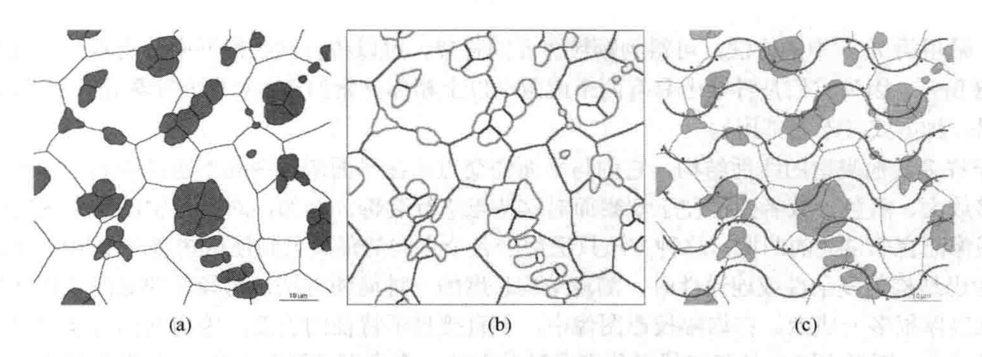


图 9.16 (a)二相显微结构的理想图形；(b)三种不同类型的相交边界；(c)对边界与摆线网格交点计数的结果。测量值和计算结果详见文中的说明

使用图 9.16 中图像所示的放大率校准，来说明和比较上面描述的、用于体积与表面积测量的各种测量过程是有益的。统计像素的数量，可计算出面积比，因此灰色相的体积比为 17.3%。在图像上放置有 90 个点的方形网格，会产生 14 个命中，使用式（9.1）计算出体积比为 15.6 ± 4.2%。对每种类型的相交面测量出边界线的长度并应用式（9.3），可得出表 9.1 所示的结果。统计网格线与每种类型的边界的交点并应用式（9.4），可得到表面积的另一个估计值，也如表中所示。图像面积为 5453.6μm²，所用网格的总长为 614μm。

表 9.1 表面积测量的比较

边　界	长度（μm）	$S_V = (4/\pi) \times$长度/面积	网格计数	$S_V = 2P_L$
白色－白色	441.96	103.2 mm²/mm³	28	101.8 ± 19.2 mm²/mm³
灰色－白色	551.24	128.7	37	134.5 ± 22.1
灰色－灰色	123.30	28.8	8	29.1 ± 10.3

不同方法所得结果基本一致，这是令人满意的，但只有计数方法允许估计精度，由此可求出应被检查的截面数量，进而得到具有期望精度的结果。注意，大量表面积可形成一个小的三维体。

9.5 长度

长度测量通常适用于那些在某个方向上被拉长而横向宽度相对较小的结构，如组织中的神经细胞或血管、复合材料中的纤维或在金属中的位移。但也可测量边缘的总长度，如金属中多面体颗粒的边缘（注意与上面介绍的 ASTM 颗粒尺寸测量关联起来），或由表面相交所形成的任何线。

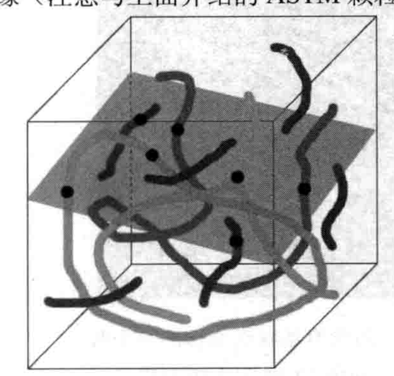


图 9.17 三维体中的线性结构图示，统计采样平面与线性结构的交点数，可测量总长度

当线性结构与采样平面相交时，结果是一个点，如图 9.17 所示。这种交点的数量与线的总长度成正比。如果三维体中的线性结构是各向同性的，或剖面相对于结构的方向是随机的，那么单位体积中的总线长 L_V（单位为微米/立方微米或长度$^{-2}$）与单位面积内的总点数 P_A（单位为数量/平方微米或长度$^{-2}$）之间的关系为

$$L_V = 2P_A$$

式中出现常数 2 的原因是，平面和线可能会以任意角度相交，就如测量表面积的情形那样。在某些情形下，可在抛光后的表面上直接统计点数。例如，复合材料中的纤维是可见的，而通过化学蚀刻可能会造成错位。

在图 9.14 所示的例子中，三个颗粒交汇的三相点表示空间中颗粒边缘的三重线。它们是扩散最快的路径。进行计数需要分

开它们，采用第 8 章中的方法，可得到颗粒边界的骨架。可以统计这种镶嵌中分支点或节点的数量，如图 9.18 所示，因为它们是骨架中具有两个或两个以上相邻元素的点。对它们计数是计算“颗粒大小”数量的另一种方式（不太常用）。

对于许多其他类型的线性结构，它们与平面的交点，在平面图像中并不容易看到，除非一些样本制备能够放大、着色，或者为了便于观察而对这些点进行装饰。例如，对硅片中的错位进行蚀刻，可生成对图像而言非常大的凹坑。这种方法只适用于各个点间的距离大到使这些点互不相接的情形。

不考虑结构的复杂性或连通性时，测量结构长度的一种简便方法是对较厚的截面（比结构的线性部分的宽度厚很多）成像。在这幅投影图像中，不直接显示特征的长度，因为它们可能向上或向下倾斜地穿过截面。图像上画出的任何线条代表穿过截面的一个表面（见图 9.19），其面积等于其长度与截面厚度的乘积。统计线性结构与图像上网格线的交点数量，然后使用式（9.7）计算总长度，就可直接测量出单位体积中的总长度。

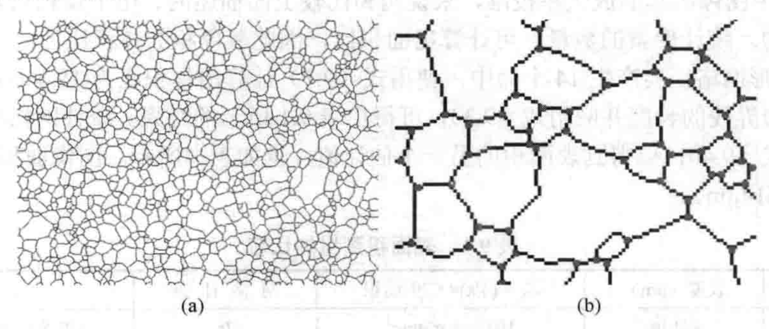


图 9.18　(a)图 9.14 所示图像经阈值处理及骨架化后的结果，骨架上有 1732 个节点(已着色)；(b)为便于观察，放大了部分结果，其中的点已被膨胀

如果采样是各向同性的，那么可使用任何网格线；如果不是各向同性的，那么如果该截面取各向同性的或随机的剖面，则可使用一个圆形网格，因为圆形网格可在平面的所有方向上采样。但在未考虑各向同性截面方向的单调性或假定样品为各向同性的危险情形下，进行测量的最实用方法是，平行于一个已知的方向切割截面（垂直切片），然后使用摆线网格，如图 9.20 所示。此时，因为它是由三维空间中各向同性的线所表示的表面法线，因此生成的摆线必须旋转 90° 到“垂直”方向。

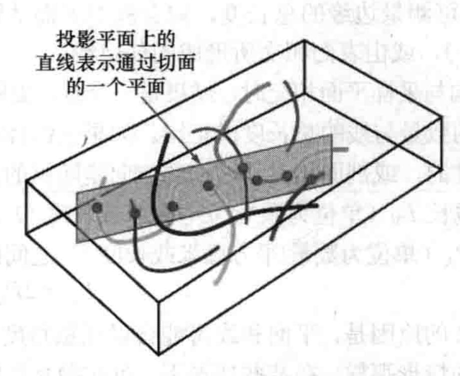


图 9.19　包含线性结构的厚切片草图。在投影图像上画一条直线并计算该直线与结构的交点数，对应于计算结构与通过切面的平面的相交数

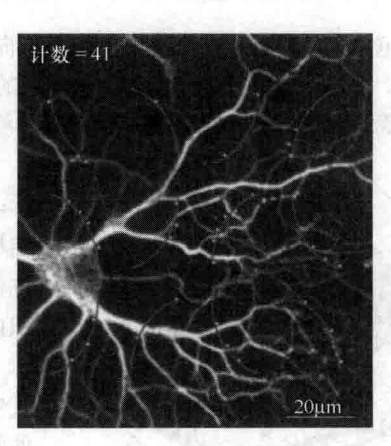


图 9.20　对含有丝状毛细管的一个截面的 TEM 图像应用摆线网格。统计与网格的交点数，可计算出单位体积的总长

在图 9.20 所示例子中，使用式（9.7）由网格线与毛细管投影的交点，可计算出单位体积的总长。共有 41 个标记的交点；对图像进行阈值处理、对毛细管骨架化，然后与网格进行“与”运算，也可统计出这些交点的数量，如本章前面及第 8 章中所示。对于面积为 12293μm² 的图像，假设的截面厚度为 3μm，网格线的总长度为 331.2μm，由式（9.7）计算出

$$L_V = 2 \times 41/(12293 \times 3)\mu m/\mu m^3 = 2.22 mm/mm^3$$

使用相同的过程也可测量出树干的总长度。

很难找到一个真正具有各向同性的三维结构。如果结构是（或可能是）各向异性的，就需要生成各向同性的截面。人们已找到了这样做的过程，但它们通常很乏味，会造成样本材料的浪费，且很难做到一致（即与物体外围相比，它们会对物体中心过采样）。

无论是测量表面积还是测量长度，生成各向同性样本的另一方法是，将三维材料分为很多小片，每个小片的方向如上所述都是随机的，然后形成可便于检查的截面。这也许是人们真正使用的最常用方法。使用这种随机的各向同性剖面时，没有必要再使用摆线网格；任何方便的网格均可达到目的。如果在平面中有一个首选方向，那么可使用圆形线网格来避免偏差。如果没有首选方向，那么可以使用方形网格或一组平行线。如果结构非常规则，则需要使用随机生成的线（计算机可方便地生成这些线）。图 9.21 显示几个这样的网格。许多现代体视学在测量期望的结构参数时，都会进行适当的样本切片并选择合适的网格。

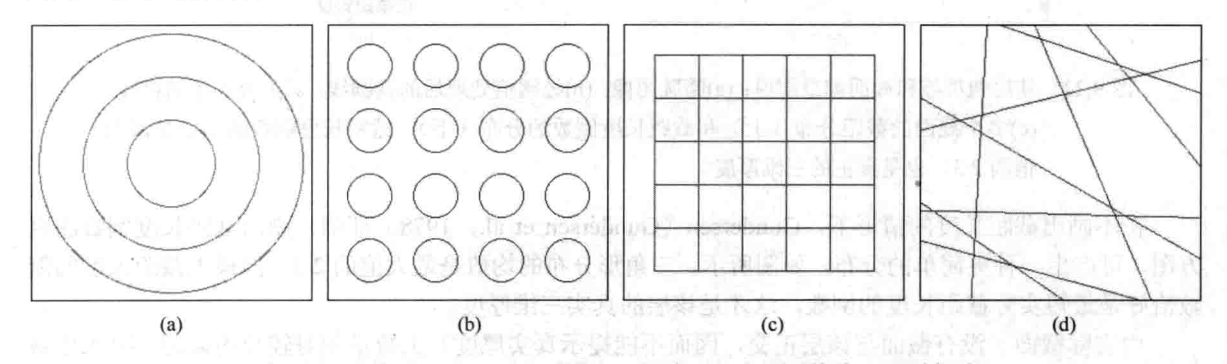


图 9.21　网格线示例：(a)同心圆；(b)圆阵；(c)方形网格线；(d)随机线

9.6　厚度

有时人们会使用平行线或径向线网格来测量层厚，如第 8 章中的图 8.16 所示。穿过该层的截面经阈值处理后，所得二值图像和这组网格线相“与”，可生成一组线段，对这组线段，我们很容易求出它们的均值和标准差。在其他情形下，当层弯曲时，测量其宽度的方法如下：对截面进行阈值处理，并使用二值图像的骨架来选取沿欧氏距离图的中线出现的那些像素，如第 8 章中的图 8.78 所示。中线值表示沿层的长度的每个点处的内切圆半径。

这两个过程都是有效和准确的，但只适用于穿过层的截面显示了实际厚度的情形。这意味着截面必须层向正交。对于表面上沉积的各层，比如图 8.16 所示的油画，这是不难实现的。但对于很多情形，如测量生物体中各种结构的厚度，就很难按照期望的方向来控制切片。层在截面中出现时才求出其宽度，是进行测量的一个经典例子，这是因为可以做到，而不是因为应该做到。测量值并不表征三维结构中层的实际厚度。截面通常要远大于实际厚度，但无法确定或校正测量中的偏差。

体视学提供了一种非常简单的解决方案，如图 9.22 所示。它首先对穿过感兴趣结构的随机方向的截面成像，然后在截面图像上画出随机线。例中的步骤如下：首先对待测层进行阈值处理，然后使用计算机生成一种随机的网格线。这些线在三维结构中的方向都是随机的。图中显示了这些线条与层相

交的长度。结果的分布表明，没有值会短于真正的三维层厚，但取决于剖面和线相对于层的方向，可能会得到更长的截距长度。注意，这一分布朝着无穷减少。

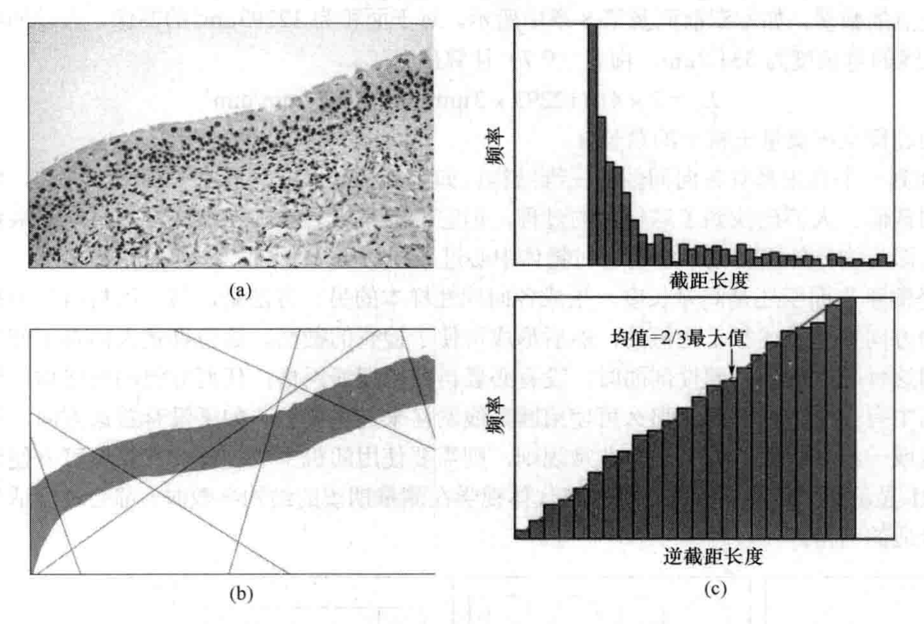


图 9.22 使用随机线和截面测量层厚：(a)膀胱图像；(b)经阈值处理后的椭圆线，其上叠加了随机线；(c)多个截面的截距分布（上）和截距长度倒数的分布（下）。截距长度倒数的均值是最大值的 2/3，也是真正的三维厚度

在不画出截距长度的情形下，Gundersen（Gundersen et al.，1978）证明，画出截距长度倒数的直方图，可产生一种更简单的分布，如图所示。三角形分布的均值是最大值的 2/3。但最大截距长度的倒数恰好是最短实际截距长度的倒数，这才是该层的真实三维厚度。

由实际截面（没有截面与该层正交，因而不能提示真实厚度）上数量相对较少的截距，可求出截距长度倒数的平均值。取截距长度倒数均值的 1.5 倍的倒数来计算真实的三维厚度，是体视学测量层厚的一种非常有效的方法，即使在任何图像中都看不到实际厚度。

9.7 采样策略

到目前为止，给出的几个例子中所使用的方法都是常用的显微方法，即通常使用光学或电子显微镜来获得通过薄切片的多个视图。这种方法的要求是，截面必须远薄于任何感兴趣的结构，但用于长度测量的厚截面除外。共焦光学显微镜不要求对薄片进行物理切片，但会产生解释起来类似的多幅等效图像。MRI 和 CT 技术也会产生截面图像。

对于不透明的材料，可使用相同的策略来切片，但图像是由反射光获得的，或由扫描电子显微镜获得的，或以仅显示表面的其他成像技术获得的。许多成像技术可显示表面下方有限深度的结构，且多数制备方法（如抛光）会产生一些表面起伏，即较软的结构往往低于较硬的结构。同样，其标准是信息深度（即扫描电子显微镜中发射电子的深度，或抛光起伏的深度）要远小于任何感兴趣结构的尺寸。对上面测量体积、面积和长度的公式进行校正，可得到适合于有限截面厚度或表面起伏情形的公式，但复杂性大增并要求精确地求出深度或厚度。

尽管给出了获得各向同性采样的过程，但目前仍很少提及如何才能得到均匀和随机的结构采样。当然，先将整个标本切割成许多小片，然后通过盲随机采样来选取其中的一些小片，最后在随机方向

的平面中切割每个小片，可得到理想的结果。但这不是最有效的方法（Gundersen & Jensen，1987；Cruz-Orive，1993；Gundersen et al.，1999；Gundersen 2002）。可以设计一个系统的或结构的随机采样，来使用最少的样本和最少的工作量产生所期望的无偏结果。这种方法适用于各个层次的采样策略，例如从动物群体中选取一些测试动物（或从培养皿中选取几个细胞等），选取几块组织，选取旋转角度进行垂直切片，选取微切片进行查看，选取成像区域，并放置网格进行测量。其他先进技术通过对采样进行加权，提高了分析的效率（Gardi et al.，2008）。

图 9.23 说明了使用这种技术从一堆苹果中选取某些苹果的基本原理。将一个篮子中的苹果数量（30）除以所需的样本数（5），就得到了用于控制样本间隔的值（6）。生成一个从 1 到 6 的随机数来控制采样的起始点。此后选取每一个序号为 6 的苹果。这就实现了对一篮苹果的均匀且随机的采样。对下一篮苹果生成一个新的随机数。结果是每个苹果有同等被选中的概率，但实际的选择可最有效地对整个群体进行采样。

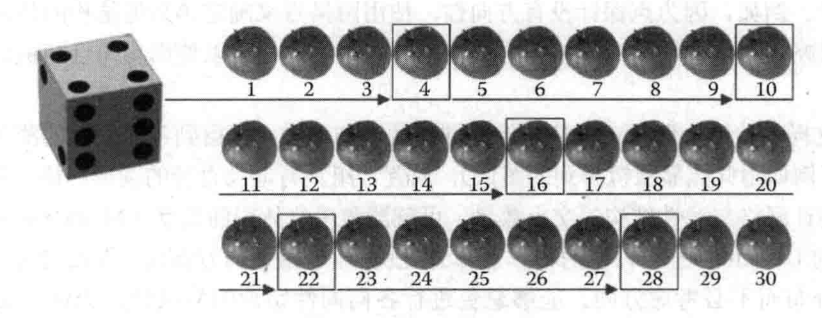


图 9.23　文中所述的系统或结构随机采样示例

使用这种技术测量图像时，为说明其过程，我们考虑求人体中骨骼的体积比问题。显然，此时不能用某个截面来代表这样一个多元化的结构：通过头部的一个截面与通过腰部的一个截面相比，能成比例地显示出更多的骨骼。对于体积比，不需要具有各向同性的截面，因此可以使用横截面。实际中使用对人体无伤害的 CT 或 MRI 扫描仪来得到这些横截面。如果截面是随机放置的，那么会看不到身体的某些部分，而有些平面会紧靠在一起而对其他区域过采样。最有效的采样会均匀地隔开平面，但必须避免经常碰到某个结构而避开其他结构的布局偏差。

假设对人体使用 8 个截面的初步采样就可达到某个期望的精度。通过生成一个随机数来进行系统和随机的采样，可将第一个截面放在人体顶部的某个八分之一处。对于身高为 160cm 的人，这个位置是顶部 20cm 内的某个位置。然后使用 20cm 的均匀间定位其他截面，使得每个截面在人体高度的 1/8 范围内都有相同的偏移。这些截面就构成了一个系统且随机的采样。对下一个人体使用一个不同的随机数，使所得截面移动一个单位。这个过程保证了每个人体的每部分均有被测量的同等机会，而这就是均匀随机采样的要求。

对每幅图像应用网格时，生成了两个随机数，以指定网络的第一个点的位置（即左上角）。然后以均匀的间隔将其他网格点移动一个单位。对显微镜载玻片上选取的成像位置应用相同的步骤。要得到足够的统计采样，如果决定从载玻片上的 8 个视场来获取图像，则将该区域分成 8 个相等的小区域，如图 9.24 所示，生成两个随机数来给出第一个矩形中的第一个视场的位置，并在每个其他区域中的同一相对位置获取后续图像。

对于垂直剖面的旋转，如果决定使用 5 个方向，如上面的图 9.10 所示，那么对于第一次切割，使用一个随机数来选择 0° 和 72° 之间的一个角度，然后以 72° 的间隔系统地放置其他的方向。这种方法显然可以推广到任何情况。它是均匀的和随机的，因为选中采样群体中每部分的概率相同。可用于探测结构的所有截面和线条，被用到的概率都相同。

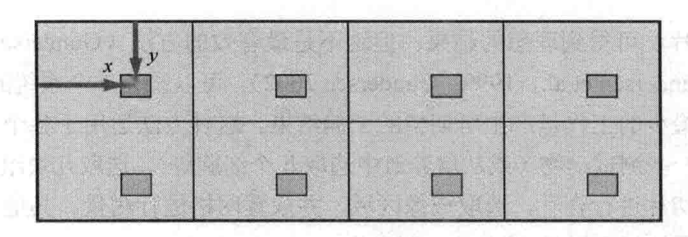


图 9.24 用于选择幻灯片上视场的系统随机采样。将该范围按将要获取的图像数目分为许多区域。两个随机数用来选择第一个区域中的第一幅图像的位置（橙色）。后续图像（绿色）接下来在每个区域同样的相对位置获取

在设计实验时，体视学者的目标是实现结构的 IUR（即各向同性的、均匀的、随机的）采样。对于某类探针，因为探针本身没有方向性，因此对标本中的任何各向异性是不敏感的，进而可以放宽对各向同性的要求。例如，因为点探针没有方向性，使用网格点来确定感兴趣结构的体积比并不要求上述的垂直切片策略。对结构均匀和随机采样的任何一组平面，均可以使用均匀且随机放置的网格来得到无偏的结果。

也会出现这样的情形，即对三维体成像并创建适当的网格，可得到各向同性的测量。对于一系列平行平面图像（例如由共焦显微镜得到的图像），放置一组具有适当直径的圆形网格，就在空间中形成了一个球体，统计网格与线性结构的交点数量，可就测量单位体积的长度（Mouton et al.，2002)。

下面描述的 Disector 是一种体积探针，因此也没有与之相关的方向性。剖面对可以穿过结构（均匀且随机地）分布而不必考虑方向。能够避免进行各向同性切片的复杂性，无疑是实验设计的一个优势。

但如上面提到的那样，对于长度或表面积测量，探针（表面和线）有方向，因此必须各向同性地、均匀地和随机地放置它们。此时需要采用进行垂直切片并使用摆线网格的策略。相比于完全随机的采样，使用一种系统随机采样方法可降低需要被检查的此类截面的数量，但它并未降低确保所有方向都有同等概率被选中的需求。它具有重复性，除非可以证明样品本身是各向同性的（以及均匀的和随机的），然后除非采用一种合适的 IUR 采样策略，否则结果将会出现偏差，且偏差大小不能确定或补偿，也不能将偏差假定为常数，进而对不同的实验结果进行比较。

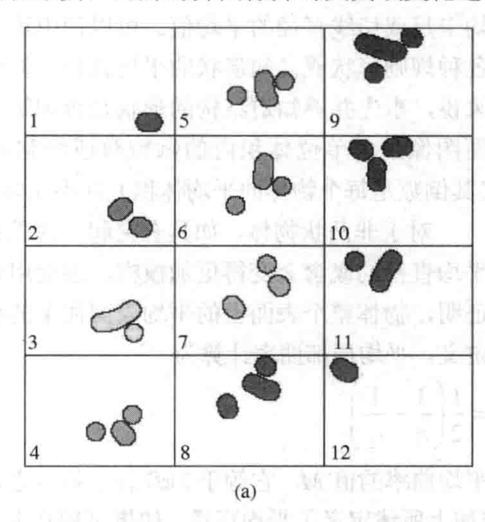
9.8 确定数量

图 9.25(a)显示了穿过三维结构的几个截面。在截面必须是 IUR 的这一条件下，由相交面或使用与网格相交的点数，可直接求出结构的总体积［见式（9.1)］。由边界线的长度，或使用边界和合适网格的截距数量，可求出结构的总表面积［见式（9.3）和式（9.4)］。由物体与截面相交所产生的单独的特征数量，可求出管状结构的长度［见式（9.7)］。但在各个截面中不能确定它是一个物体还是多个物体，如果它是一个物体，却不能确定管道是否分岔。当然，我们并不知道管道已打成结这一事实。只有通过在靠得很近的截面间内插出其他截面，进而将各个截面上的信息关联起来，并重建三维物体，才能评估这些拓扑的性质［见图 9.25(b)］。

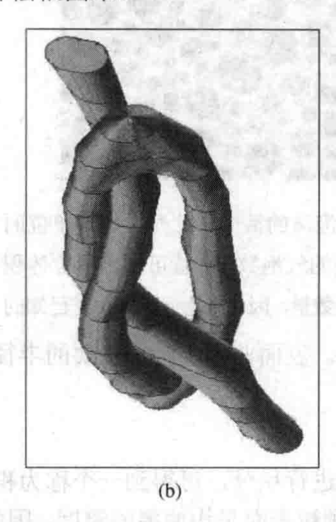
体积、表面积和长度都有可测性质。让截面穿过结构，并在截面上放置表示探针的网格线和一些点，可提供计算这些可测性质的数值信息，这些信息须满足测量精度要求，以及均匀、随机和（对面积和长度测量）各向同性采样的要求。数字的拓扑性质和连通性不能使用平面、线或点探针进行测量，这时需要直接测量样品的体积。

当然，在这种限制条件下，可对整个结构完成体成像。在某些情形下可使用无损技术，如共焦光

学显微镜、医学 CT、MRI 或声波扫描，或穿过透明体的体视观察。与计算机图形学技术结合时，它们会产生生动的视觉效果（详见第 13 章和第 14 章）。但这些方法相当耗时且成本较高，因此不适用于许多类型的样本。连续切片方法（这种方法须切割许多连续的薄片，或制备并检查许多抛光的连续平面）甚至更为困难和昂贵，因为对齐图像并对失真或变化进行补偿很困难。



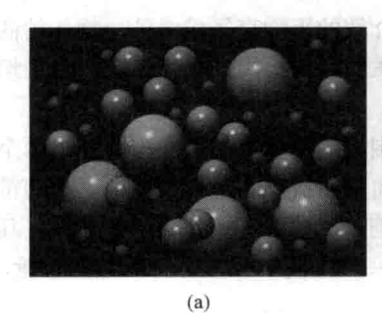
(a)



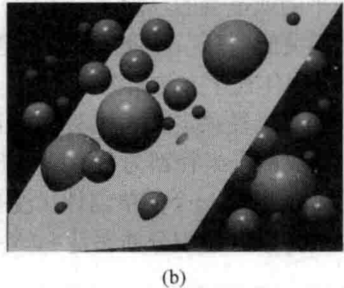
(b)

图 9.25　多个截面：(a)通过一个管状结构，可求出体积、表面积和长度。只有按序排列、对齐它们，才能重建这个三维结构；(b)排列后识别出的右手结

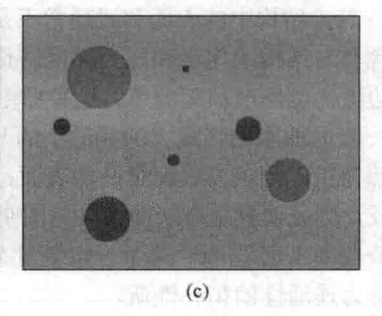
统计截面上的可见特征数量并试图由此推断三维体中物体的数量，是一种常见的错误。单位错误这一事实是出现问题的早期暗示。单位面积内成像的特征数量并不对应于单位体积内的数量。截面更有可能穿过大物体而非小物体，因此物体的大小会强烈影响与截面相交的数量，如图 9.26 所示。



(a)



(b)



(c)

图 9.26　切割三维体：(a)包含大物体（绿色）、中物体（橙色）和小物体（紫色）的三维体；(b)任意切割三维体的截面；(c)截面中出现的物体表明相交面的大小不同于物体的大小，与每个物体相交的数量并不对应于单位体积内的数量，因为大物体最可能与一个截面相关

基于特征的平均直径概念，单位面积内的数量与单位体积内的数量之间存在某种关系。对于球体，平均直径有一个简单的含义。单位体积内的数量 N_V 可由单位面积内的数量 N_A 求出：

$$N_V = \frac{N_A}{D_{平均}} \tag{9.8}$$

更确切地说，期望值 N_A 是 N_V 和 $D_{平均}$ 的乘积。图 9.27 说明了可应用这种方法的一种情形。这幅图像是穿过泡沫的一个截面。气泡的大小已知，且受空气压力和喷嘴大小控制。如果不知道气泡的大小，但知道气泡的大小一致，那么使用观察到的具有最大直径的气泡（它对应于靠近赤道的切割），可得到较

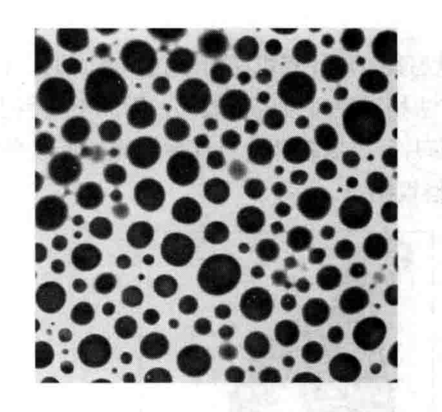


图 9.27　穿过泡沫的截面。此时，统计单位面积内的气泡数量，就可求出单位体积内的数量，因此平均气泡大小是已知的

好的估计。统计出单位面积内的气泡数量，就可由式(9.8)求出单位体积内的气泡数量。下面给出对接触边缘的特征进行计数并校正的技术。

对于非球形的凸状物体，平均直径是所有方向上平均卡尺或切线直径的平均值。可以使用这一定义来计算各种规则形状或已知形状的平均直径，但对于多数结构来说，事先并不知道结构的形状是否规则。因此，由截面图像来求单位体积内的颗粒数或细胞数或细胞核数（其倒数是每个物体的平均体积）并不实际。

对于非凸状物体，如具有突起、内部孔洞的物体，平均直径的概念会变得更加模糊，因而用处不大。业已证明，物体整个表面上的平均表面曲率的积分与物体的平均直径有关。表面曲率由两个主要的半径来定义。平均局部曲率计算为

$$H=\frac{1}{2}\left(\frac{1}{r_1}-\frac{1}{r_2}\right) \tag{9.9}$$

在整个表面上进行积分，可得到一个称为积分平均曲率的值 M，它等于 $2\pi D_{平均}$。换言之，截面上的特征数量 N_A 是颗粒表面平均曲率的测度，因此就如上所述定义了平均直径。如果可独立求出单位体积内的数量，例如使用下面的 Disector 方法，那么由这一数量和 N_A 就可计算出平均直径，进而用平均直径来表征物体。

9.9　曲率、连通性和 Disector

三维结构中的表面和线通常不是平的或笔直的，除了与上面介绍的平均直径相关外，物体的曲率还含有与结构或化学和压力梯度演化相关的信息。对处于平衡状态的颗粒结构，接触表面非常平坦，但边缘仍会出现局部曲率。这种曲率可由体视学方法来测量。

表面曲率由（最大的和最小的）两个半径来定义。从表面所围成的物体内部来观察，如果两个半径都为正，则认为表面是凸状表面；如果两个半径都为负，则表面是凹状表面；如果两个半径的符号相反，则表面具有鞍状曲率（见图 9.28）。围绕一个简单物体（没有洞、桥等）的任何封闭曲面，其总积分曲率永远是 4π。对于一组更复杂的形状，总积分曲率是 $4\pi(N-C)$，其中 N 是分开物体的数量，C 是称为连通性的拓扑性质。

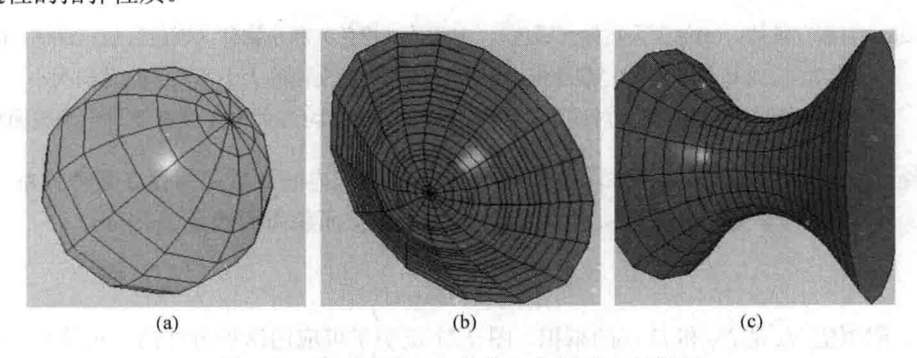


图 9.28　表面曲率：(a)凸状；(b)凹状；(c)鞍状

对于一组分开的物体，不管它们属何种形状，C 都是零（即它们不相连），且该结构的总积分曲率

可给出所出现物体的数量。单个平面表面不足以测量出曲率。此时，使用一个平面扫描探针移过样品（物理上可通过将共焦显微镜的焦平面移过样品或使用医疗成像设备实现），并统计平面探针与感兴趣表面相切的次数，可测量出总曲率。独立于鞍状相切（T^{+-}），对凸状和凹状切点（分别为T^{++}和T^{--}）分别计数。净相切数为$T^{++}+T^{--}-T^{+-}$，并得出结构的欧拉特性量（$N^{--}-C^{+-}$）为

$$(N-C)=\frac{1}{2}(T^{++}+T^{--}-T^{+-})=\frac{1}{2}T_{净相切} \tag{9.10}$$

在两种极端情形下这非常有用。其一是有多个分离（具有任意复杂形状的）物体的情形，此时 C 为零而 N 是净相切线数量的一半。这是求所测三维体内物体数量的一种简单方法。其二是只有单个物体的情形，单个物体通常是由许多扩展分支组成的复杂网络，如神经细胞、毛细血管、复合材料中所用的加固织物以及砂岩中的孔隙等。此时，N 为 1，且可测量连通性。连通性是用来将网络分开的最小切割数。对于网络内的流动或透射，它是阻止流动或透射所要求的替代路径数量，或最小障碍数量。

在许多情形下，如不透明矩阵情形，扫描相切平面并不实际。可揭示结构重要拓扑性质的最简体积探针是 Disector（Sterio，1984；Geuna，2005）。它由两个平行的检测平面组成，平面可以是透射检测的两个连续或间隔紧密的薄片。对许多材料样品而言，通过抛光成像和记录的检查平面可得到第一个平面，而向下抛光（通常使用坚硬压痕或划痕来测量距离，而且可使得图像对齐很容易）一小段距离可得到第二个平面。

Disector 逻辑比较两个平行平面中的图像，忽略持续穿过两个平面的任何特征，即使特征的位置和形状在两个平面上稍有不同。当然，两个平面必须足够近时，才可能明确地做出这个决定。通常，这意味着两个平面之间的距离必须明显小于感兴趣特征或结构的任何特征尺寸，因此与一个平面相交的多数特征也会与另一个平面相交。图 9.29 显示了各种事件的可能组合。

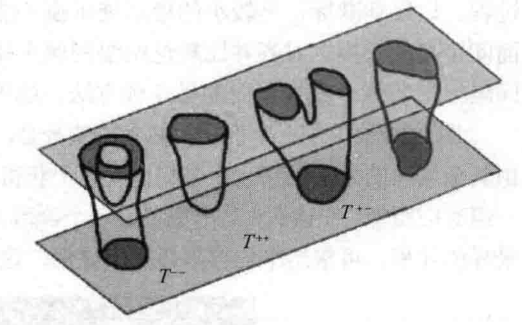


图 9.29　两个 Disector 平面中特征的各种组合。忽略那些持续穿透且无拓扑变化的特征。特征端点计数为凸（T^{++}）事件，空端点计数为凹（T^{--}）事件。分支表示存在具有 T^{+-}曲率的鞍状表面

出现在一个平面上的特征与出现在另一个平面上的特征不匹配时，则表明该三维结构在两个平面之间的某处有一个端点。这表示一个凸曲率并计数为一个 T^{++} 事件。对于一组凸状的离散物体，只出现这些端点。很少出现这样的情形，即一个平面上出现的某个中空特征在另一个平面上显示为实心特征。这表示一个空的端点，或更普遍地表示一个具有凹曲率的空洞，它被计数为T^{--}事件。结构中的任何分支（某个平面上的一个特征可识别为连接到了另一个平面上的两个特征）表示一种鞍状表面并计数为T^{+-}事件。采用用式（9.10）所示的同样方法，可对由 Disector 方法得到的相切事件求和，进而求出净相切数和结构的欧拉特性。

因为 Disector 图像中出现的多数特征会继续穿过两个截面而不提供任何信息，因此需要检查许多截面对，当然，必须对齐这些截面对来比较特征的相交。当矩阵透明且在共焦显微镜中观察时，检查和比较平行截面很简单。第 8 章中的图 8.29 给出了一个例子。使用传统的连续截面时，使用计算机处理来对齐并比较图像是有吸引力的。查找特征并忽略那些持续穿过两个截面的特征时，第 8 章中描述的标记选取（特征“与”运算）逻辑非常有用。也有必要对一个大的图像区域进行采样，这样才能使采样分布满足随机、均匀的标准。由于 Disector 是一个体积探针，因此不必关心各向同性问题。

图 9.30 给出了一个例子，其中的截面是对糖果中的气泡采用非破坏性成像方法得到的。每个圆形

图像的面积为 1.54cm²，截面之间的间距为 6.3μm。对这些图像进行阈值处理并应用布尔逻辑后，统计得到了 65 个端点。应用式（9.10）可计算出每立方毫米体积内有 3.35 个气泡。由于从图像中也可算出出气泡的体积比（6.1%），因此可将气泡的平均体积计算为 0.0018mm³（相当于直径为 0.15mm）。

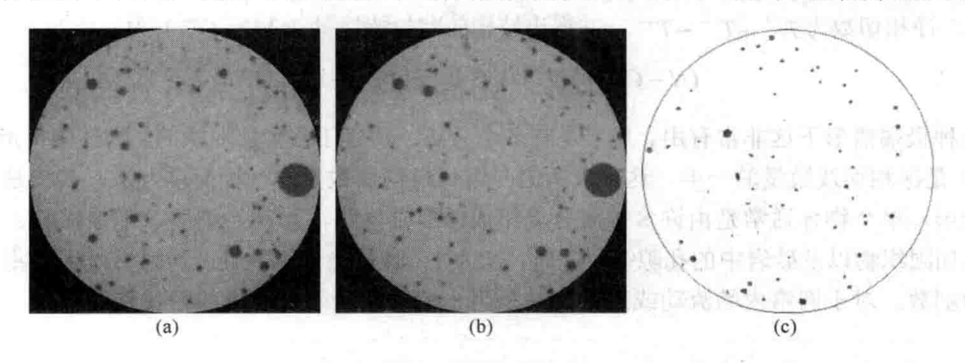


图 9.30 (a)～(b)由计算机断层扫描成像的穿过糖果的两个截面；(c)使用第 8 章中图 8.29 所示的相同过程得到的相交面

对于必须进行物理切片的不透明矩阵，问题更加严重。一般来说，需要制备一个抛光的截面，对其成像，然后在图像上放置一些用做参考的基准标记，再对样品向下抛光得更深一些，并重复前面的过程。如果基准标记是较小的锥形硬压痕（常用于测量材料的硬度），那么尺寸的减少就提供了两个平面间距离的测度。对齐并比较这两幅图像来找到新特征很乏味。表面聚焦离子束（FIB）加工，是制备间隔很密的两个成像平面的另一种方法，这种方法非常有用。

图 9.31 给出了一个例子。标本是钛合金，我们所感兴趣的是各个魏氏金属丝群落的尺寸。每个群落由其金属丝的方向识别。在金相抛光面上获得一组图像后，进一步对材料向下抛光 5μm，得到了额外的一组对应图像。在该例中，出现了另一个群落（红色箭头所指）。由此类正相切线事件的总数和由 Disector 采样的体积，可求出平均群落体积。显然，这是一个非常耗时且困难的处理。

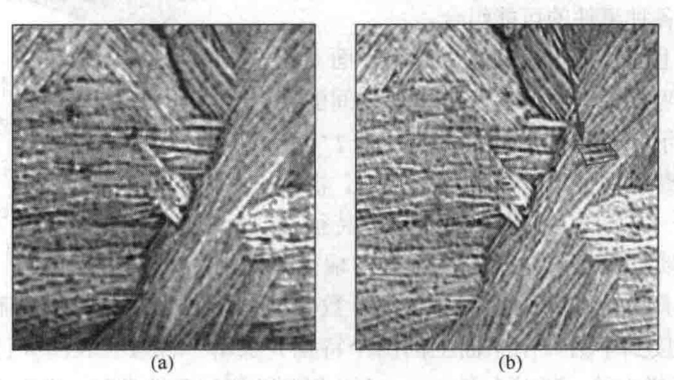


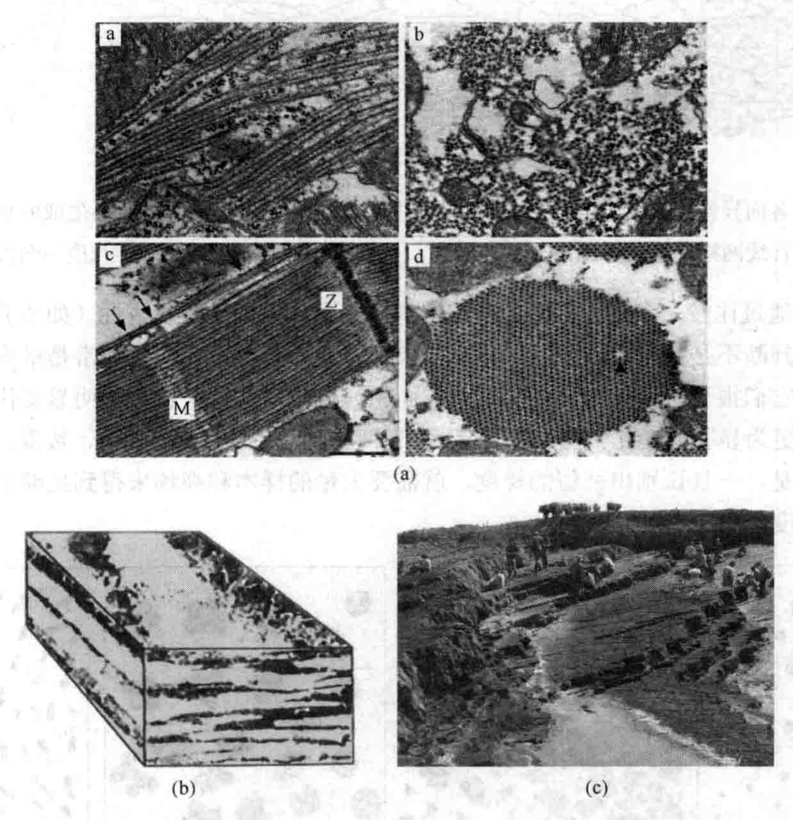
图 9.31 钛合金中金属丝群落的扫描电子显微图。这两幅图像是间距为 5μm 的平行平面。图(b)中出现了一个新群落（由红色箭头和框线指出），产生了一个正相切计数（T^{++}）

9.10 各向异性和梯度

上面特别强调了对可能不是各向同性的、均匀的和随机的（IUR）那些结构，进行切片并应用网格来消除结构中的偏差的方法。当然，当标本满足这些标准时，可以使用任何采样方法和网格，但在多数情形下，不可能假设物体是一种 IUR 结构。因此，常使用诸如垂直切片和摆线网格这样的技术，

因为不管标本是否是 IUR 的，只要在制备和测量中多付出一些努力，就可得到无偏的结果。

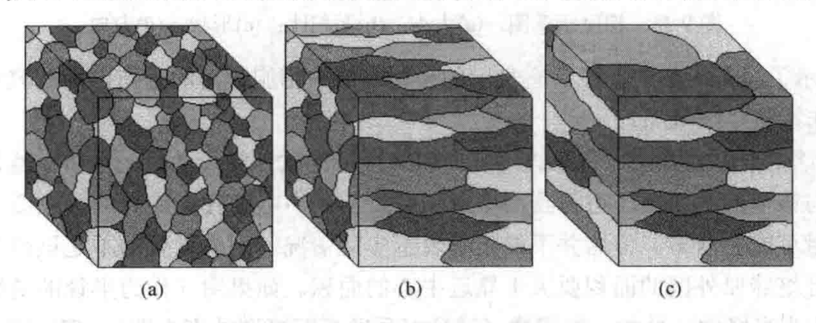
在某些情形下，测量标本中的各向异性（首选方向）或表征结构中出现的不均匀性（梯度）非常重要。这就需要付出更多的努力。首先，事先确定感兴趣的是什么类型的各向异性或梯度非常重要。有时基于物理情形可事先知道这一点。例如，材料制造过程中产生的形变和植物的生长，通常会在某个方向上产生被拉长的颗粒或细胞。沉积岩层的倾角和走向可由它的一些外露面求出。许多物理材料和生物组织在表面和外边界附近都有显著变化的结构，因此可以预计该方向的梯度。各向异性可在视觉上定性地求出，例如检查在正交方向切割的平面图像，如图 9.32 所示。通常需要检查至少两个平面来区分各向异性的性质，如图 9.33 所示。



(a)

(b) (c)

图 9.32 各向异性：(a)肌肉组织，可以看出纵向截面（a 和 c）和横向截面（b 和 d）中结构的不同；(b)对三个正交的抛光表面进行显微镜检查揭示的轧钢；(c)由岩层走向和倾角表征的沉积岩



(a) (b) (c)

图 9.33 有必要检查至少两个正交表面来揭示各向异性的性质。图中，不同的三维结构会导致相同的表面外观：(a)等轴结构；(b)针状结构；(c)板状结构

当结构中存在某个期望的方向时，经典的方法是使用网格来在不同的方向进行选择性的测量。例

如，在图 9.34 所示情形下，使用旋转的平行线网格，可在不同方向测量单位线长的截距数（即测量单位体积内的表面积）。由于每组线条测量与线条方向垂直的投影表面积，结果就显示了表面方向的各向异性。画出单位线长上的相交数的倒数（称为截距长度）与方向的关系图，可生成图中所示的极化图或玫瑰图，它表征了颗粒或细胞的平均伸长度。

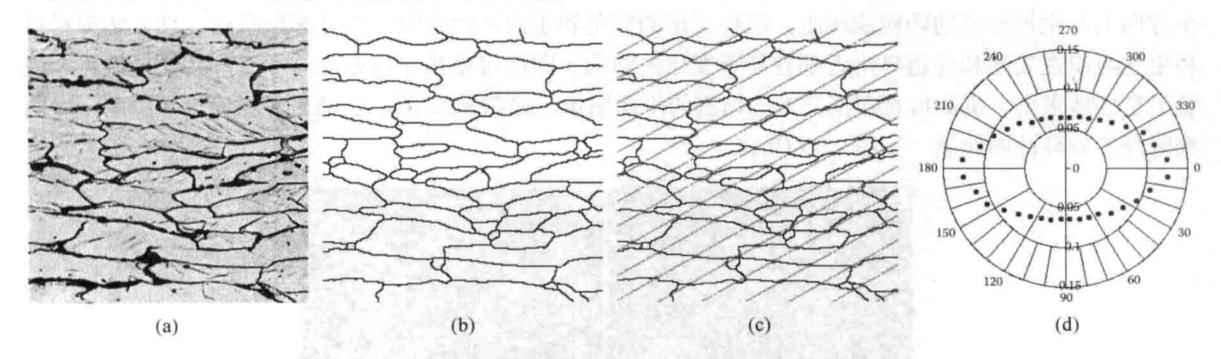


图 9.34 表征各向异性：(a)各向异性的金属颗粒结构；(b)经阈值处理和骨架化后生成的颗粒边界；(c)转动平行线网格，测量不同方向上单位线长上的截距数量；(d)平均截距长度与角度关系的极化图

梯度有时可通过比较一个标本不同部分的图像来观察，但由于度量属性（如边界线的长度和相的面积比）的视觉判断不是很可靠，所以统计地测量并比较数据来发现变化通常是最安全的。表征梯度的主要问题是，它们很少是简单的。例如，单位面积或体积内的特征数量的明显变化，可能隐藏着特征尺寸或形状的更为显著的变化。图 9.35 显示了几个图形，其中出现了大小、数量、形状和方向的梯度。第二个问题是，一旦识别出疑似的梯度，就需要大量的样本和视场来得到足够的统计精度，进而正确地表征该梯度。

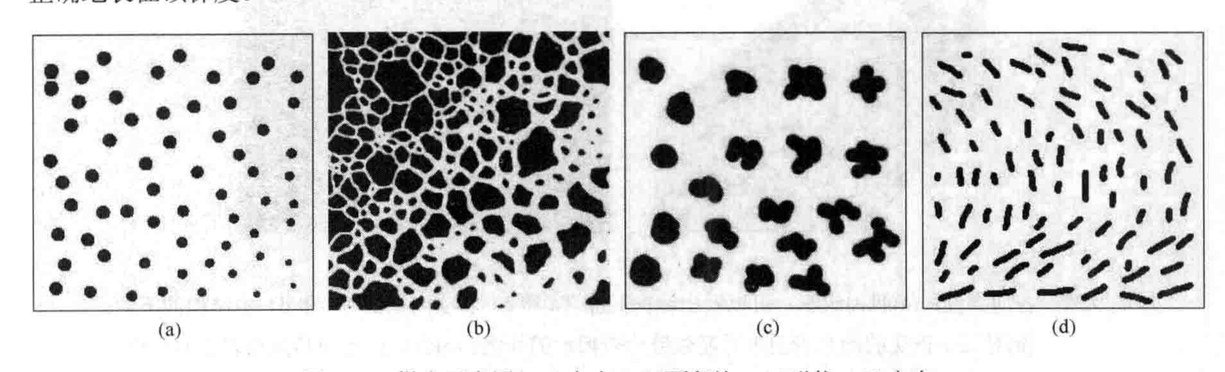
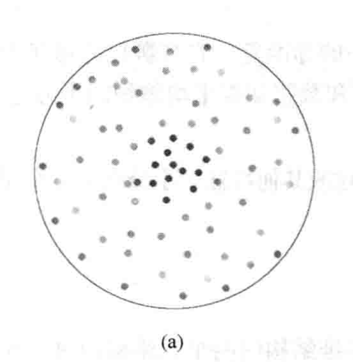
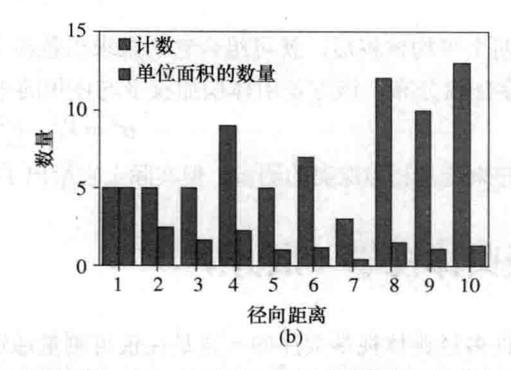


图 9.35 梯度示意图：(a)大小；(b)面积比；(c)形状；(d)方向

图 9.36 显示了一个简单的例子，它表示的是培养皿中细胞群落的分布。细胞究竟是向中心聚集还是均匀地分布在整个区域呢？

使用第 8 章中给出的方法，将来自培养皿内部区域的欧氏距离图的一个值赋给每个特征，然后画出特征数量与该值的关系图，可得到图中所示数量与到中心的径向距离的关系图。在培养皿的边缘，特征的数量可能会增多。但这并不能正确表示实际情况，因为重要的不是该数量，而是单位面积内的数量，且培养皿外围的面积要大于靠近中心的面积。如果用（作为半径的函数的）特征分布数量，除以（作为半径的函数的）像素数（这恰好是欧氏距离图的直方图），那么可得到单位面积内的真实数量变化，如图中所示，它显示了向培养皿边缘密度降低的倾向。





(a)　　　　　　　　　　(b)

图 9.36　一个区域内的特征（即培养皿中的细胞群落）分布：(a)根据到外围的距离着色的每个特征；(b)数量和单面积内数量的关系图，它是距离的函数

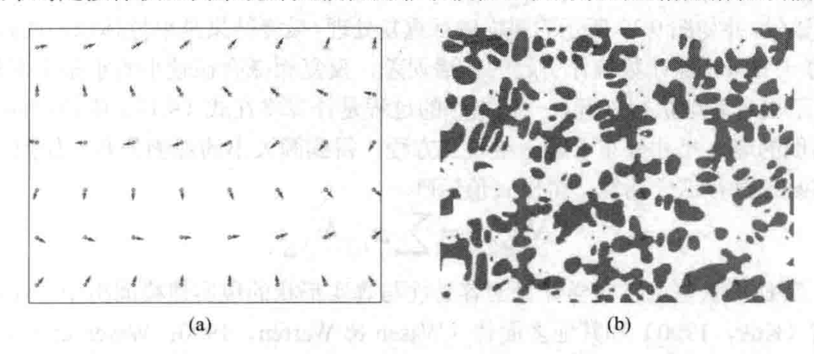
9.11　颗粒尺寸

如上所述，体视学的当前研究重点是改进采样策略，来消除非 IUR 标本情形下的偏差。二阶体视学也有一些有趣的进展，如使用两种不同类型的网格来探测一个结构，或进行两种不同且独立的测量。一个例子是使用点网格来选取特征，进行与特征体积成比例的测量（点更可能遇到大而非小的颗粒）。对于每个这样选取的颗粒，从选取点以随机方向画出一条直线，可测量该点到颗粒边界的半径。

图 9.37 给出了应用到上面所用图像中的白色树状相的过程。点网格选取了位置。对于落入感兴趣相的每个点，从每个点以均匀的随机方向画出了到特征边界的一条径向线。对于非凸形状，这条线的长度可能由必须加在一起的多条线段组成。对于各向同性的结构或各向同性的截面，线角对方向均匀采样。对于可能穿过各向异性结构的垂直截面，应使用一组正弦加权的角度。每条线的长度给出了颗粒体积的一个估计：

$$V = \frac{4}{3}\pi r^3 \tag{9.11}$$

它与颗粒形状无关。平均多个颗粒测量和半径测量的结果，可得到体积加权平均体积的一个鲁棒估计（换言之，体积加权平均是按体积比例来统计大颗粒而非小颗粒，从而求出网格点落在该颗粒上的概率）。



(a)　　　　　　　　　　(b)

图 9.37　测量采样点截距：(a)网格点阵，每个点都有与其相关的方向；(b)对图 9.3 所示图像应用该网格，并测量落入感兴趣相中的这些网格点到指定方向边界的径向距离。注意，对于非凸形状，径向距离是截距线各部分之和

定义平均体积的一种更熟悉的方法基于一个数字加权基，它等同地对每个颗粒进行计数。将相的总体积除以颗粒数，可求出平均体积。上面的例子中给出了使用点数来求总体积比的方法，以及使用 Disector 来求单位体积内的颗粒数量的方法。总体积除以单位体积内的颗粒数量，就得到了数字加权平均体积。

求出这两个平均体积后，就可组合它们来求出颗粒大小的标准差，它与颗粒的形状无关，因为对于传统的数字加权分布，该方差由体积加权平均体积的平方和数字加权平均体积的平方之差给出：

$$\sigma^2 = V_{\mathrm{V}}^2 - V_{\mathrm{N}}^2 \tag{9.12}$$

尽管看起来这是相当深奥的测量，但实际上它给出了颗粒或其他特征大小分布的一个重要特性。

9.12 经典体视学（展开）

过去，许多经典体视学关注的一直是在仅可测量通过三维结构的两个二维截面时，如何求出三维特征的颗粒尺寸（Weibel，1979；Cruz-Orive，1976，1983）。经典方法是为颗粒假设一个形状（即所有颗粒的形状相同，形状不因颗粒的大小而变化）。当然，假设的最方便形状是球体。随机切割球体产生的圆形大小分布，可按图 9.38 所示的方法计算出来。

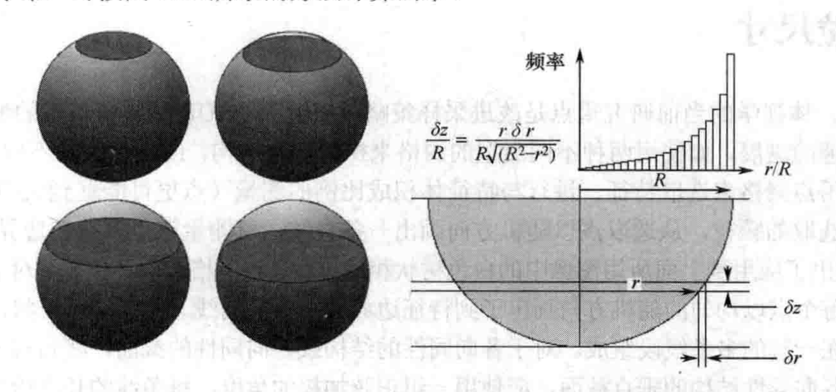


图 9.38　随机切割一个球体生成圆形尺寸的分布，它可由解析几何计算

如果所有球体的大小相同，那么知道所应生成的圆的大小，便可展开该分布（这种实现显然可以追溯至 Wicksell，1925）。在原理上（有时在实践中），该方法可按一系列步骤执行。如果将最大圆的单位面积内的数量转换成单位体积内的数目［如式（9.8）中的 $N_{\mathrm{V}} = N_{\mathrm{A}}/D$］，那么可预计出分布上每个小条中计算出的相交数，并从测量的分布中减去它们。下一个小条中保留的单位面积内的数量，可类似地转换为单位体积内的数量，并如图 9.39 所示的那样进行重复处理。最终结果是单位体积内的数量的展开分布。

这种方法的一个问题是计数统计引起的传播误差，反复相减会在较小的小条中累积误差。为防止误差累积，也为了更有效地执行展开，一种更好的过程是计算将在式（9.13）中使用的 α 值矩阵，进而为三维结构中出现的球体尺寸分布求解一组联立方程，得到圆大小的观测分布，如图 9.40 所示。电子表格程序可很容易地执行这一计算，得出 α 值矩阵。

$$N_{\mathrm{V}_{\text{球体},i}} = \sum \alpha_{i,j} N_{\mathrm{A}_{\text{圆},j}} \tag{9.13}$$

除球体外，其他形状也可在数学计算的容易性与真实形状的现实建模间折中。这些形状包括椭球体、圆盘、圆筒（Kok，1990）和其他多面体（Wasen & Warren，1990；Wasen et al.，1996），其中的一些可用来填充空间（如真实的细胞和谷粒那样）。这些形状存在 α 系数表，系数表可直接用于测量在截面上观察到的二维特征的颗粒尺寸。

但这种方法有几个严重的缺陷。首先，理想的形状通常并不存在。细胞核不是理想的球体，细胞、气泡和颗粒都不是相同的多面体形状（较大的颗粒与较小的颗粒相比，通常有更多的表面和边缘），金属中的沉淀颗粒很少是完美的圆盘状或条状，等等。即使是分开的气泡也不总是完美的球体。三维物体形状的知识对于展开法的成功至关重要，因为交点的大小分布会随形状剧烈变化（图 9.41 给出了一个立方体与球体相比较的例子）。大小随形状变化很常见，这种变化会导致结果出现偏差。形状假设方

面的微小误差会使得结果产生很大的误差。例如，球体上出现的一些不规则物或突出物，会导致许多较小的相交特征，我们很容易将它们误认为是许多小球形颗粒。

图 9.39　将最大物体单位面积内的数量转换为单位体积内的数量，减去预计的最小相交数，下移到下一个粒度级并重复该计算，展开了颗粒尺寸

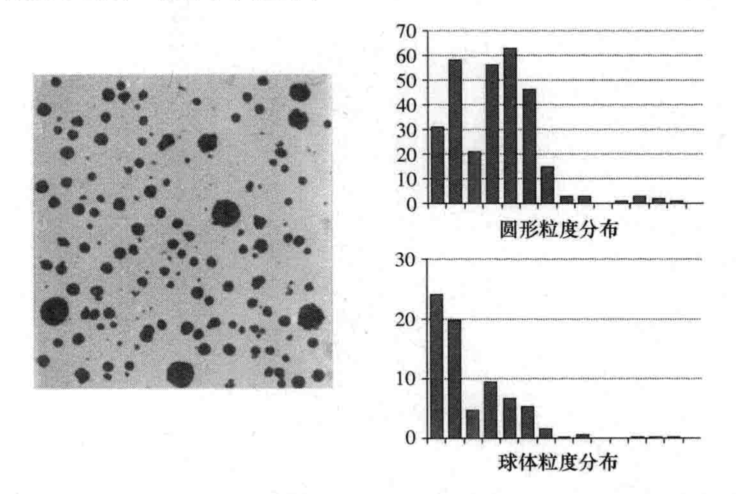


图 9.40　展开一个颗粒尺寸。样本是含有石墨球的铸铁；注意，虽然已在展开中使用了球形模型，但特征并不是完美的圆形

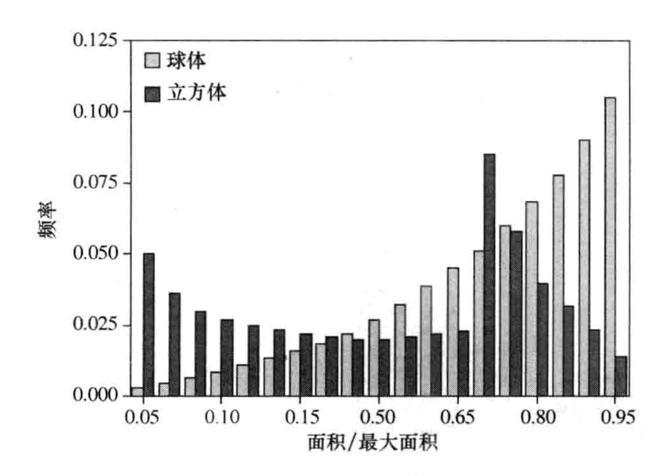


图 9.41　随机切割一个球体和一个立方体所产生的截距面积的颗粒尺寸

第二个问题是，从数学观点来看，展开计算是病态的和不稳定的。计算统计导致了二维特征数量分布与尺寸关系的不确定性。对于很少遇到的几个最大特征，变化在分布的顶端通常最大。但在展开过程中，这些误差被放大，在一些大小类别中往往导致特征的负数量，更加放大了统计的不确定性。

由于这些原因，虽然几十年来展开法一直是一种广泛使用的经典体视学方法，但在最近的 25 年中它已逐渐失宠。相反，人们已开始将重点放在使用更强大的工具来找到无偏的方法测量结构参数。但抛开这些限制不说（或许是因为它们的重要性和效果未被得到充分认可），人们仍然经常执行颗粒尺寸的展开（Ehtezazi et al.，1999；Nooijer et al.，2008）。在三维物体的形状未知时，即使所观察到的截面特征明显不是圆形，也通常假设它们是简单的球体。逐个样本进行比较所给出的一些数据是错误的，且背离了应用体视学过程的规则。

第10章　特定特征的测量

对图像中每个特征进行的测量一般可分为4类：亮度（包括颜色值和参数，比如与亮度相关的密度）、位置（绝对位置和相对于其他特征的位置）、大小和形状。在每一类中，可以进行许多不同的特定测量，同时执行这些测量的不同方法也很多。多数图像分析系统在每一类中至少提供几种测量。用户有时会发现必须从几种不同的测量参数中进行选择。问题是确定哪些测量参数是最有用的或最适合于解决特定问题的。

下面所示的对大小、位置和亮度的不同测量并不都是普通的，但至少熟悉大小、位置和亮度的基本概念。形状是一个更加复杂的主题，将在下一章中描述。

测量通常会产生适合进行统计分析、特征选取或图形显示的一个数值输出。通常情况下，数据的解译会留给单独的程序（如数据表格或专用统计或数据分析软件包）。在有些情形下，这些数字会在决策中使用，或成为分类的基础，详见第12章中的讨论。这样的例子有，零件中孔洞大小和位置的质量控制测试，根据识别出的癌细胞进行医学病理决策，或识别某个场景中的不同物体（或人）。

10.1　亮度测量

通常情形下，本书讨论的各类图像中，每个像素都记录了表示原始场景中对应点的亮度的数值。将几个这样的数值组合起来，可表示颜色信息。最典型的亮度值范围是0～255（8位范围），但也可能会遇到更大的范围，如10位或更多位，最高达16位（0～65535）或更多位，具体取决于相机、扫描仪或其他获取设备的类型。多数从输入源（如具有超过8位深度的冷却式相机和扫描仪）接收图像的程序，会将它们保存为16位图像（对于灰度图像，占用2字节/像素，对于彩色图像，红值、绿值和蓝值占用6字节/像素），即使其实际灰度或色调分辨率稍小于16位。

有些系统对所有位深使用数值范围0～255，但对具有超过8位深度的图像，报告称其使用小数值而非整数。这样，就可很容易地比较使用不同位深获得的图像，并且这里采用的约定也是如此。在极少数情况下，存储的值可能是实数（如高程数据）而非整数。但在多数情形下，这些图像仍然用一组整数值存储，因为很容易操作这些数组并将它们转换为显示。在许多情形下，需要保留一个校准表或函数，以便将像素值转换成有意义的实数。不管这些值表示的是高程、密度还是其他属性，这都适用。

为特殊成像设备或某幅特殊图像创建这样一个校准函数的过程非常重要（Inoue，1986；Chieco et al.，1994；Swing，1997；Boddeke，1998；Ortiz & O'Connell，2004）。用于获取图像的许多相机和其他设备，既不是完全线性的，也不是精确对数的，像素数值与输入信号（如光子强度）间的关系也不是完全一致的。许多相机有一个随整体光照变化而变化的输出函数。自动增益电路或用户可调伽马控制的出现，使得更难建立和维护任何一种校准。根据一个白点，使用相机内置的软件或自动软件对彩色图像进行调整，最终取决于图像主题，且除了光照的变化外，还可能受背景色或阴影的影响。通常，任何自动增益或暗电平电路、自动色彩平衡等，都将阻挠对相机的校准。对于消费级摄像机和静物相机，关闭这种“功能”并不总是可能的。

对于彩色成像，相机可能包括自动白平衡调整功能（适用于完全不同于大多数科学和技术图像的

场景）。使用消费级静物相机或视频摄像机时，几乎无法控制光源的色温，也不能保持相机响应的一致性。第 1 章中给出了使用从标准颜色图表中测量的 R、G、B 亮度来校正颜色数值的几个例子。要特别注意的是，典型彩色相机具有对应于相对较宽范围波长的 R、G、B 传感器，且实际波长和亮度的许多不同组合可以产生相同的 R、G、B 值。因此，对此类相机几乎无法执行颜色测量，加上光源、观察角度等的影响，问题会变得更加复杂。第 1 章中的色彩校正过程产生的是视觉色彩匹配，而不是光谱光度测量。

即使可以保证使用稳定的光源和一致的相机设置，仍会存在亮度校准问题。有些标准可以合理地得到，如阶梯状胶片形式的密度标准。测量这样一个标准区域的亮度时，通常需要使系统处于校准状态。在扫描大面积样本（如电泳凝胶或 X 光胶片）时，可将一些密度标准集成到每幅扫描图像中，以便可直接进行校准。在其他情形下，可以周期性地引入各种标准样本来检查校准。

光密度定义为

$$\mathrm{OD} = -\lg\left(I / I_0\right) \tag{10.1}$$

式中，I / I_0 是穿透样本而未被吸收或散射的入射光部分。如果相机/扫描仪及其光源已经过精心调整，使得线性亮度值的整个范围涵盖从 0.1（未曝光胶片灰雾级的一个典型值）到 2.5（中等密度曝光的胶片）的光密度范围，那么得到的亮度和光密度间的关系如图 10.1 所示。曲线形状是对数的。

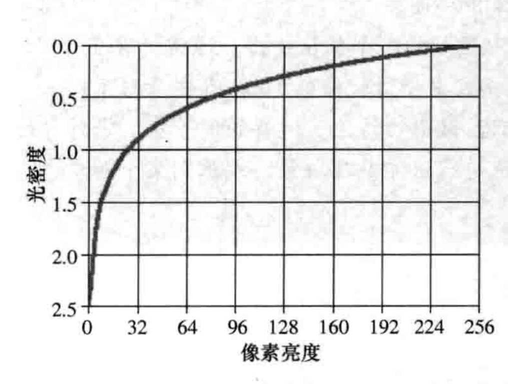


图 10.1　光密度与像素亮度间的关系，其中像素亮度已被调整到线性范围 0.1（灰雾级）～2.5（典型的曝光胶片）

在相对较低的密度/亮度像素值处，敏感性相当好。刻度亮端处的 1 像素亮度值差异，对应于小于 0.01 的光密度。但在高密度/暗像素刻度端，1 像素亮度值差异对应于 0.3 的光学密度变化，这个变化非常大。它表明，在试图应用线性相机或扫描仪来测量光密度时，需要超过 8 位的灰度。此时，可结合使用精度更高的输入设备和查找表（查找表可将值转换为对数光学密度），然后将结果存储为 8 位图像。

当遇到更高的密度值时，这个问题会变得更加严重。含有大量银颗粒的胶片，如 X 光胶片或 TEM 胶片，其密度可达 3.5～4.0。数字化此类胶片时，若要得到令人满意的结果，就需要使用至少具有 12 位/通道精度的扫描仪，最好是 14 位/通道精度的扫描仪。人们总是喜欢扫描底片而非正片，但有时无法得到底片而必须使用正片，尽管正片没有较大范围的亮度值。Ortiz and O’Connell（2004）中讨论了动态范围的非线性和损耗。

为有机分子的一维或二维分离目的，在待扫描胶片中包含一个灰度刻度板是很常见的。图 10.2 显示了一个例子。已知密度值允许为扫描仪构建一个校准刻度板，然后使用该刻度板来测量每个物体的光密度。在这一应用中，每个斑点中的蛋白质总数与整体光密度成正比，整体光密度是平均密度与面积的乘积。适用于显微镜的光密度标准源很少，彩色标准源更少，但有时可通过染色来制备一些彩色标准源（Hoffmann et al.，2005）。

使用校准刻度板会引起另一个问题。图像中的特征通常包含有一些像素亮度变化。平均密度（或根据像素亮度校准的其他量）通常不是线性记录的。在求密度的平均值、总量或其他校准量时，考虑到这一点很重要。如果只对特征内的所有像素进行亮度平均，然后使用校准刻度转换该值，那么会得到错误的结果。此时，先将每个像素的值转换为校准值，然后对它们求和或平均，才是正确的。不包含那些不是特征一部分的相邻背景像素同样也很重要，这意味着特征轮廓的勾画要准确，而不能是任意形状（如圆或方形）。手工描出的轮廓通常不可接受，因为手工描出的轮廓通常会大于特征

的边界。解决办法是使用前几章中的程序来对图像进行阈值处理并将图像分割成一些离散的特征。二值图像中的像素定义了特征，且可用做掩模来定义灰度数组中的像素。

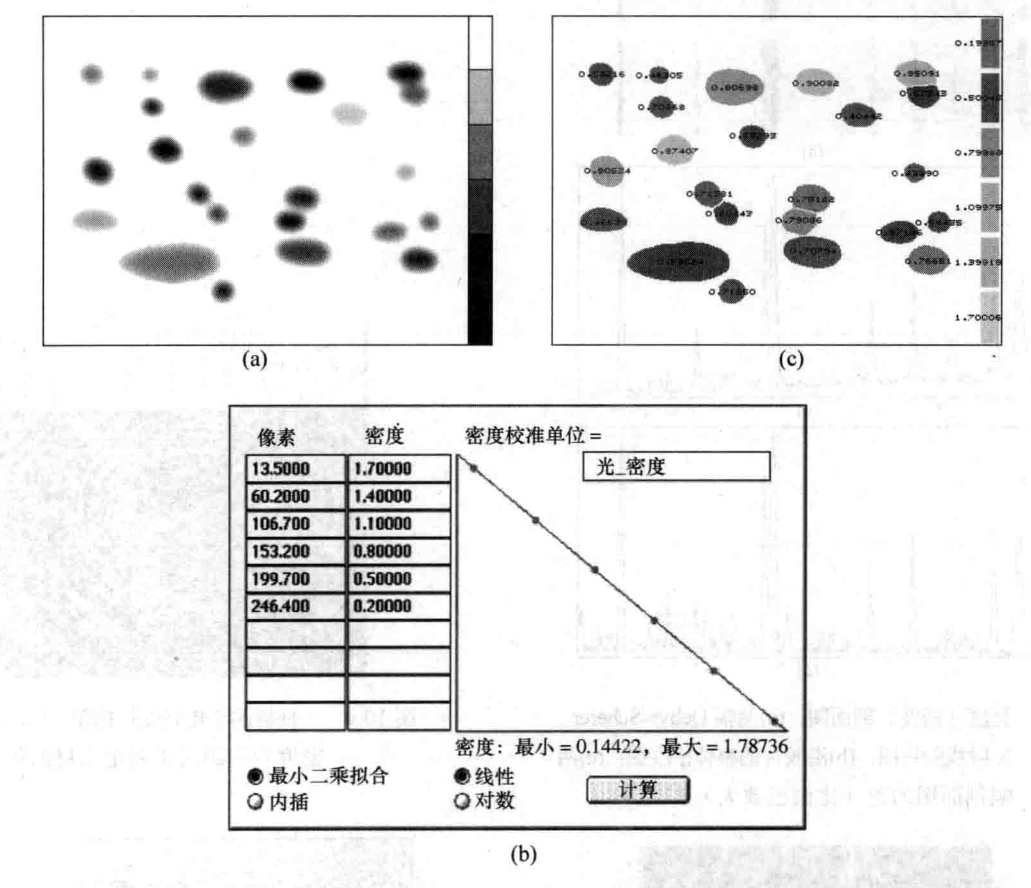


图 10.2　二维电泳分离的扫描图像：(a)原始图像，其中包含了一个灰度校准刻度板；(b)图(a)中灰度阶梯板的校准图，它显示了这种特殊扫描仪的光密度和像素亮度间的关系；(c)通过使用校准图计算每个点的平均光密度，由图(a)得到的测量数据

对于密度的一维测量，过程是相同的。图 10.3 显示了在 Debye-Scherer X 射线相机中曝光胶片的一个例子。垂直线出现在与目标材料的晶格结构中的原子间距有关的位置（角度），亮度上的小差异和其他线的出现表明存在微量的其他化合物。在已知线条垂向扩展的情形下，通过平均垂直方向上的值可降低噪声，并沿胶片得到更精确的密度变化图。在对像素值求和前，正确的步骤是将各个像素值转换为密度，这样才能得到相对线密度的正确值。减去密度值的线性图可找到差异，或结合频谱处理理程序可检测尖峰。类似的过程适用于测量大量样本，包括本质上属于一维变化的植物年轮（见图 10.4）和湖泊沉积的年融积层。因为感兴趣的是峰和谷的位置与间距，因此有些像素不要求进行亮度校准。木材内部出现的结构（细胞壁）可能会加大图形的变化，对于测量目的，这种变化可认为是噪声。对图像应用高斯平滑改善了剖面图（见图 10.5）。

图 10.6 显示了电泳分离中架或柱的亮度轮廓。由于厚度或光照的不均匀性，使用在架间测量的亮度剖面可简单地校正背景变化。亮度值跨过各柱的中心进行平均，避开了边缘。柱间各条带位置的校准，通常使用放置在相邻柱中的已知分子重量标记来完成。这是必要的，因为在这些样本中的距离尺度通常不是线性的或一致的。

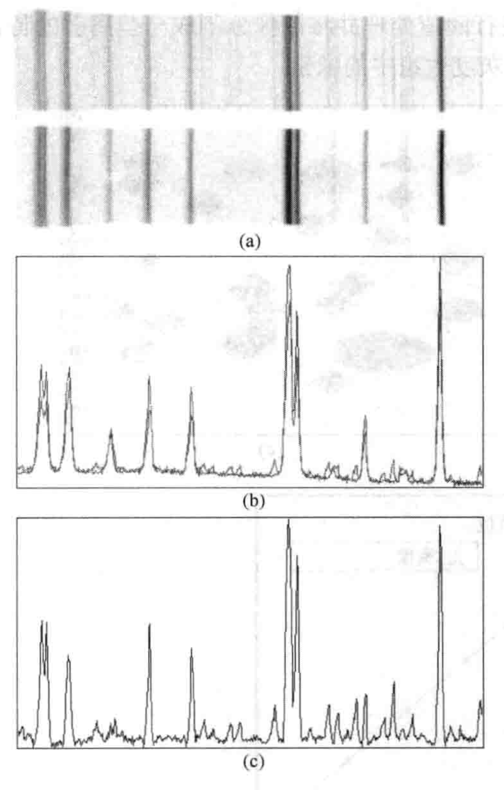


图 10.3　亮度（密度）剖面图：(a)两幅 Debye-Scherer X 射线胶片图；(b)沿胶片的整体密度图；(c)两幅剖面图的差（比例已放大）

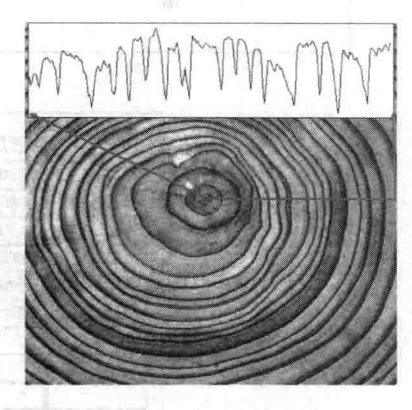


图 10.4　一棵树的年轮亮度扫描图。年轮的密度和间距用于测定木材的年代

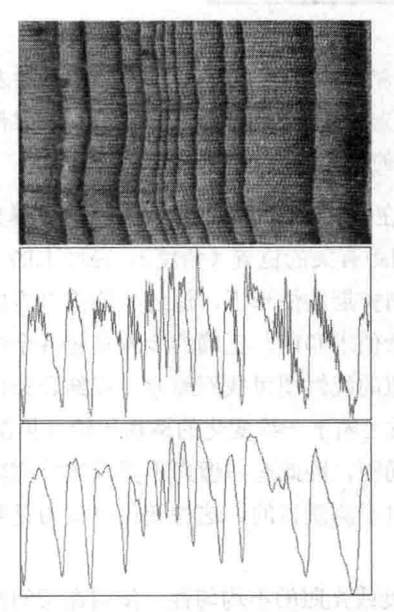


图 10.5　木材的高倍放大图像，显示了各个细胞和年轮。上图显示了图像的亮度剖面。在下图中，图像已被平滑来降低各个细胞的影响，并更好地表征了年轮

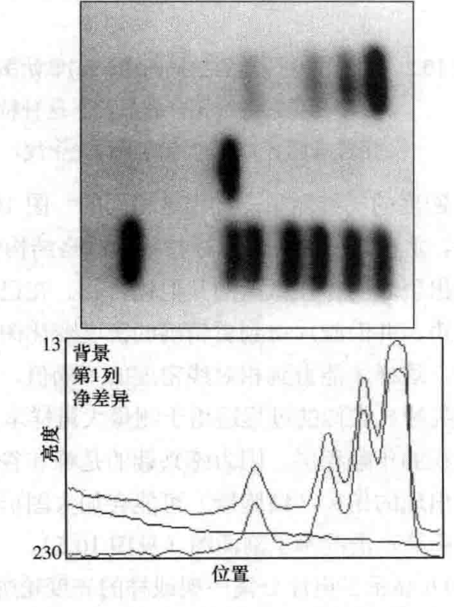
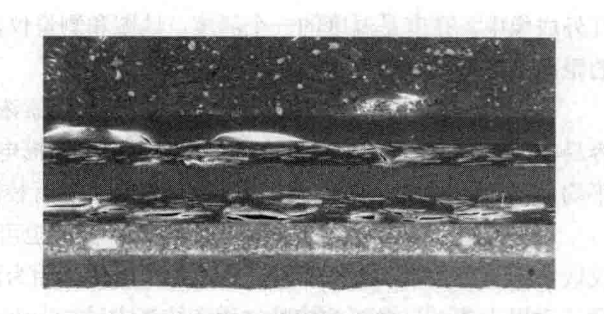


图 10.6　一维蛋白质分离。通过减去轨迹间的扫描，调匀了沿每个轨迹中心的亮度扫描（密度剖面图对应于顶部轨迹）

剖面图也是应用网格线来测量涂层厚度的一种有效替代方法。图 10.7 显示了漆层的横截面，以及沿垂直方向平均像素值（水平方向平均）的变化图。不同剖面显示了可用于求平均厚度的瞬变，图中每个瞬变区域的宽度可用于求厚度的变化。色相值图具有最陡的瞬变，并最好地定义了层厚。



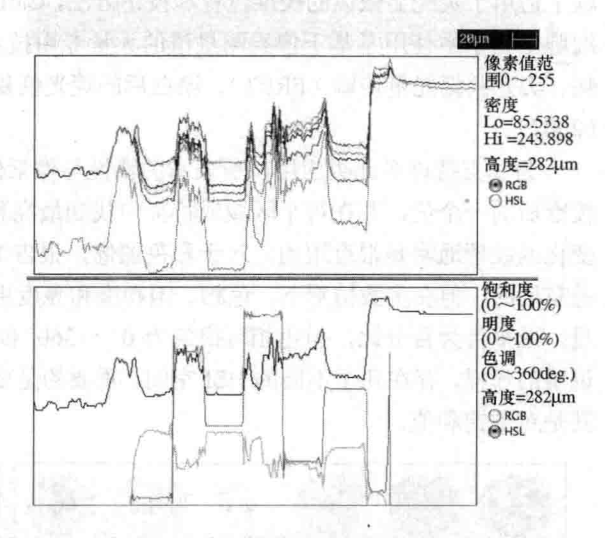


图 10.7　穿透多层油漆横截面及剖面数据。图轴是从上到下的垂直位置。上图显示了红、绿和蓝亮度，下图显示了色相、饱和度和亮度分量

并非所有图像的记录像素亮度和正被成像样本的性质（如密度）间都具有方便的关系。在显微图像中，成分或结构差异也会产生对比度，还会出现衍射或偏振效应。在真实图像的表面中，局部表面方向和纹理会影响亮度，就像光源的颜色和表面颜色间的相互影响那样。卫星影像必须去掉云层和表面覆盖（比如积雪）。在多数此类情形下，亮度测量并不提供关于样本的信息。我们的目标通常是采用相对的通道亮度值来区分将被组合在一起的区域，或分离这些区域。

数字化彩色图像，拥有足够位数来为每个彩色平面提供足够精度的问题会被放大。许多平板扫描仪可为每个 RGB 通道获取超过 8 位的图像，通常至少为 12 位，使用它们可得到颜色信息，然后可创建一幅优化的 8 位图像，进而发送到计算机中。即使这提高了精度，也很难实现精确的颜色表示。如第 1 章中指出的那样，与视频摄像机或数字静物相机相比，彩色图像数字化时，需要更好地控制相机设置和光照，除非存储时使用了 RAW 模式，并仔细考虑了光照的特性。即使是提供了更好照明一致性且通常使用具有滤波器的线性 CCD 阵列来得到 RGB 信息的平板扫描仪，也不是测量实际颜色信息的较好选择。滤波器会选择较宽范围的波长，且完全不同的亮度和颜色组合可产生相同的结果。但这样的扫描仪校准后，可以在计算机显示器或打印输出上重现足够的颜色。

校准通常是通过扫描或拍摄一个色卡来完成的，如第 1 章中图 1.48 所示的 GretagMacbeth ColorChecker 色卡。自 1976 年以来人们就一直使用这个图表（McCamy et al.，1976），它最初是摄影、电视和印刷的测试。目前的各种颜色测试和颜色配置中，仍在制作并使用它进行测试。除了灰度亮度和三原色外，选择了 12 个色块来覆盖色彩空间的中间位置，特别是在现实摄影中肤色（特别是在广告中，此时产品和印刷目录之间的颜色匹配非常关键）很重要的情形下。每个色块的观测值用于构建相机或扫描仪的校准曲线，故可以实现一致性。用来构造和应用这些曲线的软件通常带有颜色检查图表。

要特别强调的是，这些校准曲线适用于视觉色彩匹配而非彩色测量。分光光度计测量亮度高谱和滤长的关系。数字化图像只记录跨越（由 RGB 滤波器遍历的）一个较宽范围波长的整体亮度。

当然，像素亮度值不必是光密度或颜色分量信息。图像是人类进行沟通的有效信息，因此人们几乎会为各类数据使用图像。卫星图像会记录多种多样的信息，如云量、云顶的温度、海面温度等。在

红外成像中，亮度是温度的一个测度。地图和测量仪器都将表面高程显示为图像（第 15 章将介绍表面的距离图像）。

在来自 SEM 的 X 射线图像中，亮度值与元素浓度近似成正比。这些关系不一定是线性的，也不容易校准。X 射线发射强度受其他元素影响。扫描电子显微镜的反向散射电子图像的亮度值，会随着平均原子数的增加而增大，所以根据已知标准进行校准，也可求出样本上各个小区域的化学成分。

即使是在最常见的成像思想表示中，像素值也可能与染色浓度及标本中的污渍有关。荧光亮度不仅取决于染色技术和组织特性，也取决于时间，因为漂白是一种常见的现象。Zwier et al.（2004）中回顾了适用于荧光显微镜的校准过程和校正方法，Carlsson（1990）中特别介绍了针对共焦显微镜的校准。细胞生物学中使用了基于像素测量值的大量专用技术，如荧光强度与荧光寿命成像（FLIM）、发射损耗、荧光共振能量传输（FRET）、漂白后的荧光恢复（FRAP）、荧光相关光谱学（FCS）等，详见 Cox（2007）。

为了支持许多此类应用，所表示的特性与像素值之间需要一些校准。例如，通过测量平均亮度或校准后的一个值，并在每个区域或特征中找到最亮和/或最暗像素值，或找到亮度值的标准差，来测量变化或纹理通常是很有用的。对于彩色图像，报告 R、G、B 分量（通常是范围 0～255 内的值）可能是有用的，但在多数情形下，色相、饱和度和亮度更为有用。饱和度和亮度通常在 0～255 的刻度上测量，或报告为百分比，但色相可报告为 0°～360°间的一个角度，如图 10.8 所示。注意，如第 1 章中说明的那样，存在几个不同的 HSI 空间，重要的是要知道用哪一个空间来转换存储的 RGB 像素值，尤其是对于饱和值。

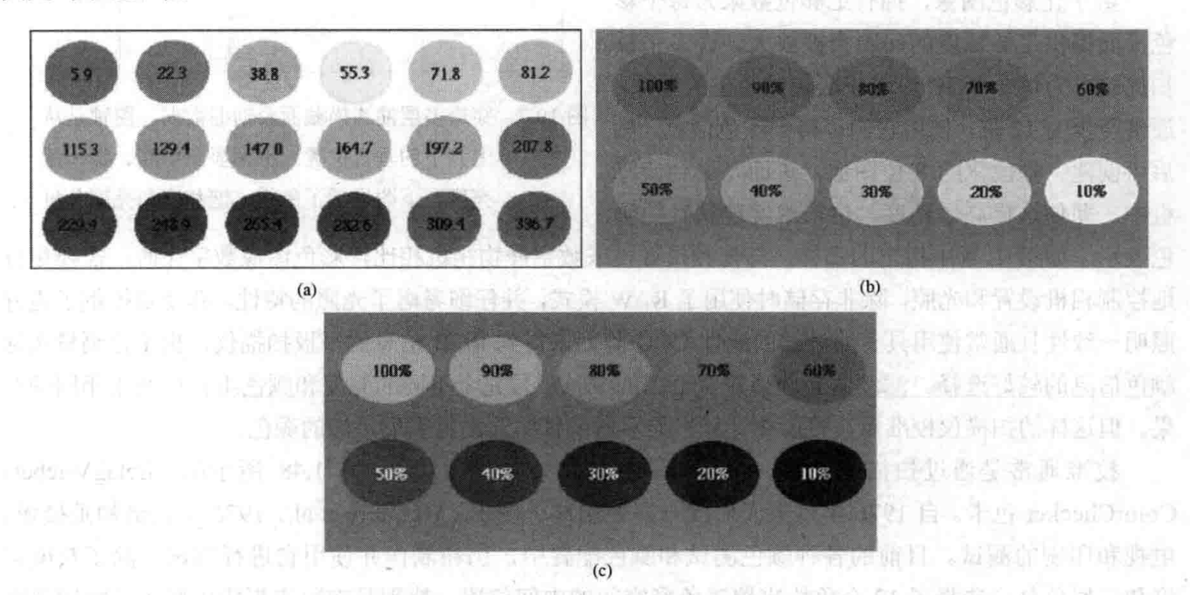


图 10.8　一系列色点的平均颜色值：(a)具有整个饱和度的特征的色相，它标记为 0°～360°间的角度；(b)具有常数色相和亮度的特征的饱和度，它标记为百分比；(c)具有常数色相和饱和度的特征的亮度，它标记为百分比

这些例子充其量只是可能使用像素值的少数样本。但存储值的测量和至某个校准尺度的转换是一种广泛使用的技术。数据的统计分析提供了均值和标准差、图像内或图像间的位置比较等。对于这样的工作过程，重要的是建立有用的校准曲线，这样做需要关于样品的标准和/或基本知识，并物理上创建和记录图像，而这完全超出了本书的范围。

10.2 确定位置

在涉及测量亮度值的图 10.2 至图 10.7 所示的例子中，为了对结果进行解译，也需要用到特征的位置。对于跨越许多像素的不规则特征，位置的定义有多种，其中的有些定义更容易计算。例如，一个特征的中点的 x, y 坐标，可确定为组成特征的像素的最小值和最大值之间的半途。通常情况下，像素地址本身只是位置的整数计数，多数常从数组左上角开始。这一约定源于多数计算机显示器工作的方式，即从左上角开始使用光栅扫描。可能存在一些全局坐标系，单幅图像只是这种坐标系中的一部分，且这些值也可能是将像素尺寸校准到某些真实单位（如经纬度、显微镜载玻片边缘或工件的毫米数）的整数或浮点数值。

通常很难为像素尺寸建立真实的校准，且在移动标本或相机（例如，通过移动显微镜载物台或卫星）的同时保持这一校准会出现各种问题。在体育电视转播中，生成叠加到橄榄球场上的首次进攻线路和其他图形的技术，要求仔细测量相机的位置，还要具有跟随动作转动相机方向的编码器和复杂的计算机软件。在气象图上叠加行政区划边界是非常耗时的过程，好在一年中只需要这样做几次（因为卫星位置会季节性改变）。

通常情况下，在图像本身中定位一些已知特征可以建立校准所用的标记，进而精确地定位图像中的其他特征。通过移动样品、载物台或相机，使得连续观看图像部分重叠，可将位置校准从一个视野移到下一个视野，但由于很难将每个视野中的同一特征定位到几分之一个像素，因此误差会累积并快速增大。这还是在图像中没有光学失真或扫描非线性的假设下。依赖于载物台移动来得到位置信息，取决于这种机理的精度（如果运动方向相反，那么由于反冲作用情形会变得更糟）和样本相对载物台的移动的保真度。在某些应用中，将 GPS 接收机内置到一些相机中，以便自动地在图像文件中记录位置信息，可方便成像。

找到水平和垂直方向上最大坐标和最小坐标的像素后，就很容易确定某个特征的最低限和最高限。这些受限坐标定义了该特征的一个外接矩形，然后可很容易地将该矩形的中点计算为该特征的一个位置。但中点通常并不是首选的位置表示，因为它很容易偏移几个像素（如特征其他部分伸出的触须）。但计算机绘图程序中使用了这些矩形坐标。许多此类程序允许用户选取一些绘制对象，然后自动地对齐它们。程序提供的选项通常有顶部对齐、中心对齐、底部对齐、水平左边缘对齐或右边缘对齐。它们都只是矩形坐标和中点。

有些情形下，外接矩形的中心很有用，如图 10.2 中的凝胶扫描情形，因为 x 坐标和 y 坐标对凝胶的创建方式很有意义。对于多数真实图像，情况并非如此。当 x-y 轴任意时，外接矩形是有偏差的，且对物体方向敏感，如图 10.9 所示。有些教材将“最小外接矩形”定义为涵盖特征的最小面积，但在实践中旋转外接矩形使其面积最小并不容易，并且仍假设矩形对目标的形状有一些意义。

特征更具代表性的几何中心是外接圆的中心位置（见图 10.10）。与只使用外接多边形的顶点，多次旋转坐标轴方向（通常每隔 10° 旋转一次），并找到每个方向的最大坐标限和最小坐标限来匹配物体相比，使用外接圆来匹配物体外围的点会更有效。外接圆的大小和中心位置不随轴或特征的方向变化。

对于非凸状特征或内部含有孔洞的特征，几何中心可能并不位于特征边界内。几何中心也不随实际特征的内部变化而改变。对于不规则形状的特征，考虑特征的形状及所有像素的位置更为可取。这种方法定义了特征的质心（或重心），在这个点上使用针尖就可平衡特征，因此具有唯一的 x, y 坐标。质心的坐标可通过平均物体中的每个像素的坐标来求出：

$$\mathrm{CG}_x = \frac{\sum_i x_i}{\text{面积}}, \quad \mathrm{CG}_y = \frac{\sum_i y_i}{\text{面积}} \tag{10.2}$$

式中，“面积”是像素的总数。该公式给出了不限于整数的一组坐标。物体的质心或重心可以精确到亚像素，对于定位某个场景中的特征而言这非常重要。图 10.10 比较了一个不规则特征的几何中心和质心。

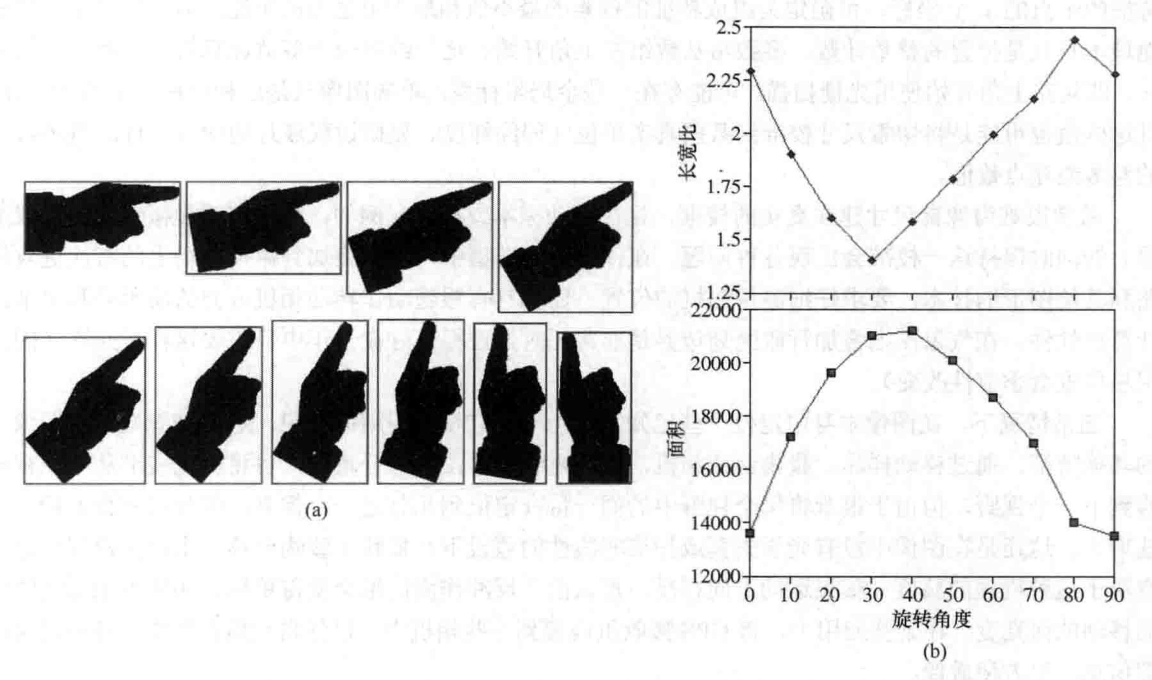


图 10.9 (a)旋转的特征和外接矩形；(b)随旋转画出的外接矩形的大小和形状变化

如果仅使用边界像素按照式（10.2）来计算质心，那么有时（尤其是在特征由链码表示时，详见第 7 章）结果会不正确。计算出的点会偏向边界中最复杂且包含有最多像素的边界部分（见图 10.11）。这种偏差也会随着边界关于像素阵列的方向变化而改变，因为方形像素在对角方向上的尺寸要大于在水平和垂直方向上的尺寸。



图 10.10 不规则特征的质心（绿色）、由像素密度值求出的密度加权质心（红色）以及由最小外接圆求出的几何中心（蓝色）

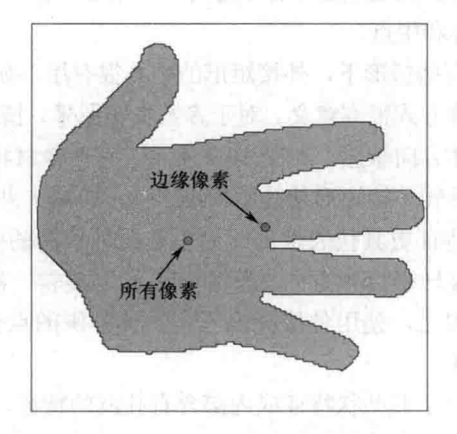


图 10.11 与使用特征中所有像素求出的正确位置（绿色）相比，只使用边界像素求出的质心位置（品红色）已向右侧偏移

由一种边界表示（如链码表示）可以正确地计算出质心的位置。正确的计算是对边界（或对多边形，或对每个顶点）上的每个点使用一对坐标 x_i, y_i，其中 x_0, y_0 和 x_n, y_n 是同一个点（即边界表示是由同一位置的两个端点来闭环的）：

$$\mathrm{CG}_x = \frac{\sum_i (x_i + x_{i-1})^2 (y_i - y_{i-1})}{面积}, \quad \mathrm{CG}_y = \frac{\sum_i (y_i + y_{i-1})^2 (x_i - x_{i-1})}{面积} \tag{10.3}$$

面积计算为

$$面积 = \frac{\sum_i (x_i + x_{i-1})(y_i - y_{i-1})}{2} \tag{10.4}$$

当使用该数据来计算像特征面积或质心这样的参数时，描述特征的链码和边界表示方法将不再具有吸引力。

式（10.2）中定义的质心或重心会平等地对待特征内的每个像素。出于某些目的，像素亮度或用校准曲线由像素亮度计算出的值，会使得某些像素比其他像素更重要。例如，在较早前的光密度计例子中，显示出的斑点和线条的精确位置就得益于这种加权。通过在式（10.5）的求和中包含这种导出的亮度值，很容易就可引入这种改进。图 10.10 比较了密度加权质心（使用像素灰度值作为密度）和平等处理所有像素的传统质心。

$$\mathrm{CG}_x = \frac{\sum_i 值_i \cdot x_i}{\sum_i 值_i}, \quad \mathrm{CG}_y = \frac{\sum_i 值_i \cdot y_i}{\sum_i 值_i} \tag{10.5}$$

分母现在是累积密度（或可能与亮度相关的任何参数）。这一计算需要访问各个像素亮度值和合适的校准曲线，因此不能与特征的边界表示一起使用。

对于一个非凸形状，质心可能并不位于特征边界内。当需要始终位于不规则形状边界内的唯一一个有代表性的点时，可由欧氏距离图（见第 8 章）确定最终的腐蚀点或 UEP，如果有多个 UEP 时，则使用具有最大值的 UEP，它是特征中最大内切圆的圆心，可用于表示位置。

10.3 方向

与特征质心位置密切相关的是如何确定其方向。此时可使用不同的参数，包括特征中最大尺寸的方向［圆周上距离最远的两个点之间的直线，也称为最大费雷特（Feret）直径或最大卡尺维度］，以及与特征边界匹配的椭圆的主轴方向。但是，就像质心可比中点更好地描述特征位置那样，由图像中所有像素定义的一个方向，往往要比任何圆周上出现或缺少一个像素时所定义的那些方向好，因为获取图像时的意外或噪声可能会轻微地改变特征的边界。

特征的转轴是特征围绕其旋转的直线，如果将它从均匀的硬纸板上剪下来，则它有最小的转矩。在各个像素到转轴距离的平方和最小的意义上，它是能最好匹配所有像素的轴。所用标准与构建图形时将线条拟合到数据点的标准相同。确定转轴及其方位角很简单，只需对图像中的所有像素坐标求和和求积。对于式（10.5）中的例子，可以使用像密度这样的值来对每个像素加权，而不是平等地对待每个像素。最方便的计算过程是将一组求和相加，如下式所示：

$$\begin{gathered} S_x = \sum x_i, \ S_y = \sum y_i \\ S_{xx} = \sum x_i^2, \ S_{yy} = \sum y_i^2, \ S_{xy} = \sum x_i y_i \end{gathered} \tag{10.6}$$

为特征累加这些求和后，x 轴和 y 轴的净矩及最小转矩角可由式（10.7）来计算：

$$M_x = S_{xx} - \frac{S_x^2}{\text{面积}},\ M_y = S_{yy} - \frac{S_y^2}{\text{面积}},\ M_{xy} = S_{xy} - \frac{S_x S_y}{\text{面积}}$$

$$\Theta = \arctan\left\{\frac{M_{xx} - M_{yy} + \sqrt{(M_{xx} - M_{yy})^2 + 4M_{xy}^2}}{2M_{xy}}\right\} \tag{10.7}$$

图 10.12 显示了测量特征方向角的一个例子。如下面解释的那样，未测量与像场边缘相关的那些特征。如图 10.13 所示，这种测量也适用于线。此时，轴突已被阈值处理并骨架化，且删除了节点，只留下了线段。为表征轴突的方向分布，画出线段总长与角度而非数量的关系图更有意义。由于这些线段不是直线，因此式（10.6）和式（10.7）由骨架中的所有像素计算出的角度，要比仅使用端点计算出的角度更有意义。图 10.14 给出了组合有位置和角度的一个例子。组织中细胞的方向随垂直位置变化。如第 9 章中讨论的那样，这是一个梯度的例子，它此时由图像内各个特征的测量值表征。

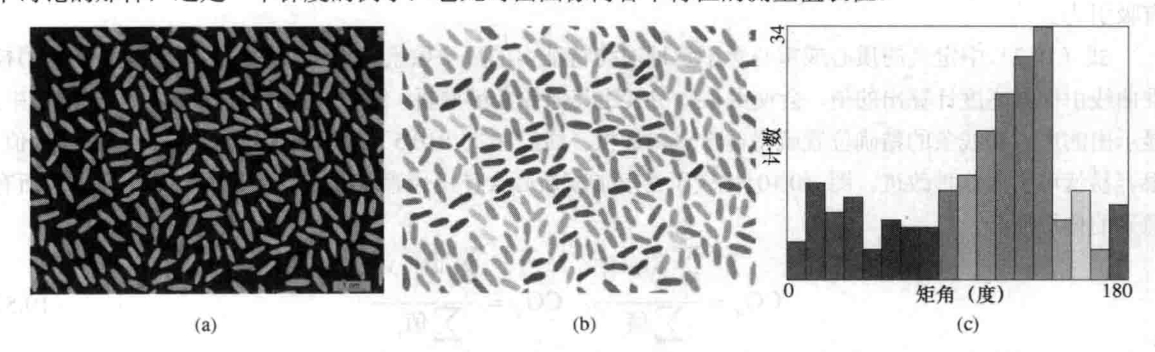


图 10.12　特征方向的测量：(a)米粒图像；(b)根据角度着色（色相）后的图像；(c)米粒的角度分布

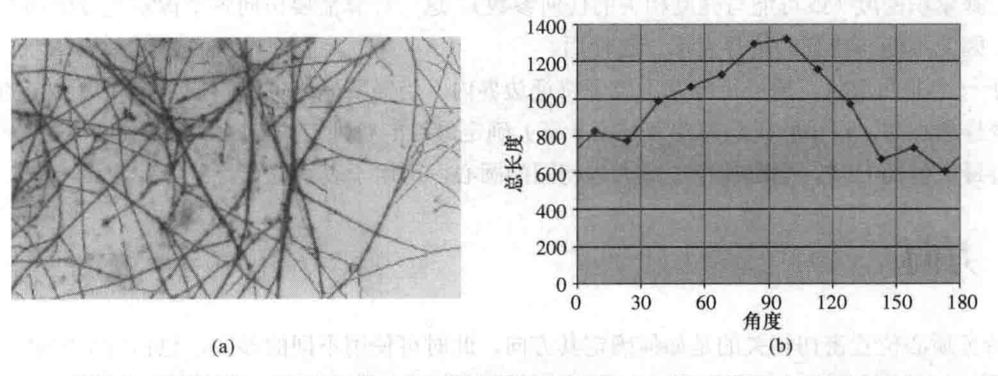


图 10.13　纤维方向的测量：(a)轴突图像，其上叠加了骨架线段（如第 8 章所示，删除了节点）；(b)纤维的累积长度与纤维角度关系图，给出了首选方向

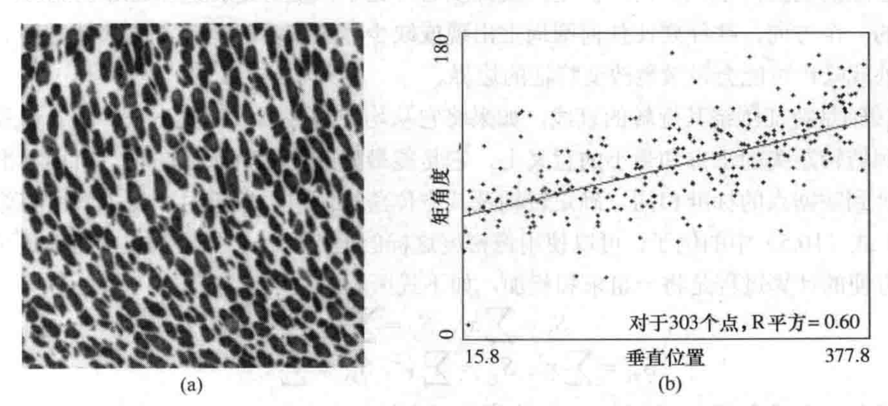


图 10.14　(a)组织中的细胞；(b)角度梯度与垂直位置关系图

10.4 相邻关系

在某些应用中，单个特征的位置坐标可能没有相邻特征之间的关系重要。例如，图 10.15 显示了特征的几种分布。如何描述这些分布才能揭示它们随机展开、隔开（自回避）或聚类的程度呢？

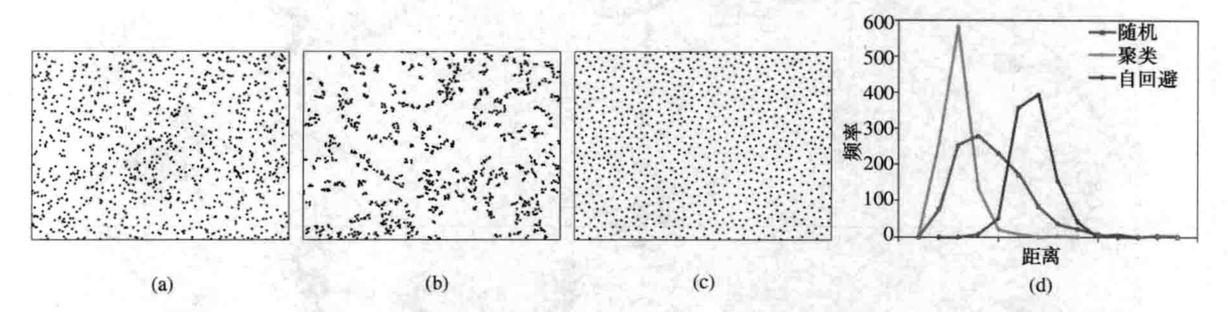


图 10.15 特征分布说明：(a)随机排列；(b)聚类排列；(c)隔开或自回避排列；(d)每种点分布的最近邻距离直方图。聚类分布的均值小于间隔分布的均值，而间隔（自回避）分布的均值大于随机分布的均值

Schwarz and Exner（1983）证明，最近邻距离分布直方图可给出一种答案。也可使用任何一对相邻特征间的距离（如第二近距离），但多数情形下最近邻域特征对是最容易辨识的。求出表示每个特征的质心点的坐标或其他坐标后，对整个结果表排序，可直接找出每个点的最近邻点，但这很耗时（最好交给计算机来完成。计算出这些点间的直线毕达哥拉斯距离，并用它构造直方图。这反过来又可以由分布的均值和方差（或标准差）来表征。这些结果可用于分析物体聚类之内和之间的分散度（Luo & Koo，2007）。

在处理与视野边缘相邻的特征点时，一定要小心：如果到边缘的距离小于到视野内最近邻的距离，则不应在分布中使用这个距离，因为它可能会引起偏差（Reed & Howard，1997）。视野之外的另一个特征可能实际上靠得更近。对于包含许多特征的大图像（视野），这个问题通常并不重要。最近邻域距离方法也可推广到三维情形；在三维成像情形下，这种方法会结合共焦显微镜使用（Baddeley et al.，1987；Russ et al.，1989；Reed et al.，1997）。也可以对所有相邻对进行统计测试（Shapiro et al.，1985）。

考虑图 10.15 所示的特定分布。图 10.15(a)所示图像是点的一种随机分布，通常称为泊松随机分布，因为最近邻域距离的直方图呈泊松分布。例如，这是将盐洒到桌上时所观察到的那种点分布。每个点与其他点完全无关，因此是“随机的”。对于这样一种分布，最近邻间的平均距离恰好是

$$\text{平均距离} = 0.5\sqrt{\frac{\text{面积}}{N}} \tag{10.8}$$

式中，N 是视野内的点数。对于泊松分布，方差等于均值。因此，对于随机分布的点，由单位面积上的点数就足以求出最近邻距离直方图的均值和方差。

当点分布中出现聚类时，多数点至少都会有一个相当靠近的邻点。因此，平均最近邻距离减小了。在多数情形下，作为聚类点之间的间隔均匀性测度的方差也会变小。如图 10.15(b)中的图像所示，该聚类会生成最近邻距离的直方图［见图 10.15(d)］，与为随机情形下获得的泊松分布相比，该直方图更窄，且均值更小。对天上星星的研究表明，星星会强烈地聚类（为星系和星系团）。有些种类的鸟和鱼会聚类，人也会聚类，即聚集在城市和乡村。

如图 10.15(c)所示，当点是自回避点时，也会影响到最近邻距离。与随机情形相比，该直方图的均值更大，如图 10.15(d)所示。作为点间间隔均匀性测度的方差，通常也会下降。自回避或规则分布通常很常见，肌肉组织中的线粒体、金属中沉淀的颗粒，都是这种分布，因为物理扩散方式对排列有重要

影响。线粒体的这种分布可为纤维提供均匀的能量。形成一个沉淀颗粒会耗尽该元素周围的基质。沙漠上仙人掌的生长过程中也会出现同一效应，即会争夺资源（如水资源）。理想情形下，为吸引客户到新市场购物，城镇购物中心的位置应该是自回避的，但现实生活中它们往往沿主干道聚类（见图 10.16）。



图 10.16　北卡罗来纳州罗利市的购物中心位置。平均最近邻距离为 0.91 英里，它小于随机分布的期望值 1.43 英里，因而表明存在聚类

在相同面积上，最近邻距离分布的均值与随机分布的均值之比，可测出特征的聚类或自回避倾向，而分布的方差则可测出这一倾向的均匀性。因为随机情形下的均值和方差仅取决于在图像区域内出现的特征数量，所以这种比较很容易进行。

使用特征的位置坐标找到最近邻对，也可表征特征分布的各向异性。与最近邻间的距离测量不同，方向是通过每个特征到最近邻来测量的。对于各向同性排列的特征，最近邻方向应是角度的单值函数。将直方图画成玫瑰图后，给出了对该单值函数的任何偏差和各向异性的程度。

例如，如果测量图 10.15(c)中规则间隔分布的那些特征点，那么玫瑰图近似为圆形，并表明该分布是各向同性的。图像在水平方向上拉长 10%而在垂直方向上收缩 10%时，单位面积的特征总数保持不

变。视觉外观图像（见图 10.17）对一位随意的观察者而言，并未表现出各向异性。但最近邻方向的玫瑰图（见图 10.18）表明多数特征现在有一个最近邻特征，该特征位于上方或下方，而不是均匀地分布在各个方向。玫瑰图的形状是这种类型各向异性的一个灵敏指示器。

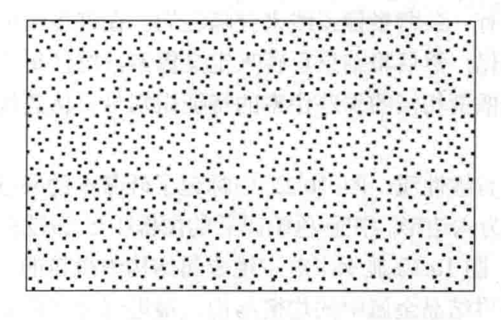


图 10.17　水平拉伸并垂直压缩图 10.15 (c)中的点分布，在最近邻方向上引入了不均匀性

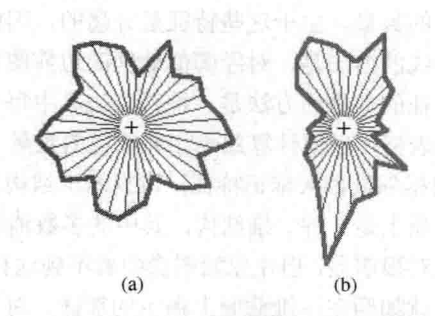


图 10.18　特征数量与最近邻方向间的关系图（玫瑰图）：(a)图 10.15(c)；(b)图 10.17

当特征尺寸与特征间的距离相比较小时，可使用特征质心来测量最近邻距离。当特征尺寸大于特征间的距离时，尤其是特征的形状不规则、大小变化且具有不同方向时，那么测量边缘之间的距离而非质心之间的距离可能更适合。Yang et al.（2001）表明，这些距离的方差分布对不均匀和各向异性聚类最为敏感。最小间距是某个特征与其最近邻特征之间的边缘间的最短距离。如图 10.19 所示，它可能完全不同于质心间的距离，且实际上可能会涉及不同的相邻特征。

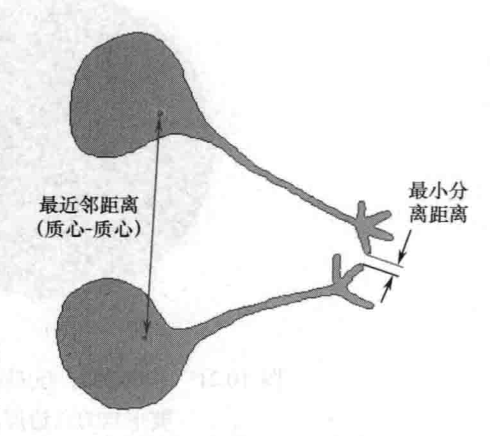
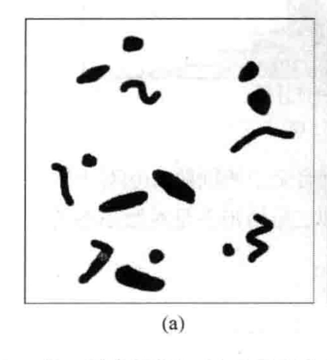
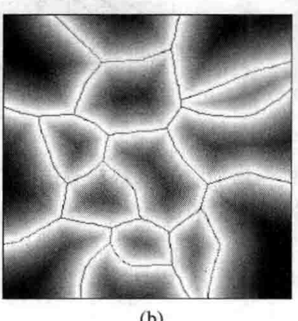


图 10.19　两个特征的质心间距和最小间距

在第 8 章中介绍的欧氏距离图（EDM）可用于求边缘间的距离，如图 10.20 所示。特征之间背景的 Skiz 骨网或骨架是一幅由泰森多边形（Voronoi）曲面组成的图像，其中每个多边形中有一个特征。这种曲面单元的 EDM 测量每个像素到 Skiz 骨网上线条的距离。将这些 EDM 值赋给特征（见第 8 章中使用的掩模运算），可得出每个特征到 EDM 的一个最小像素值，它是到具有最接近边界点的特征的距离的一半，因为 Skiz 骨网从两个特征的正中间穿过。





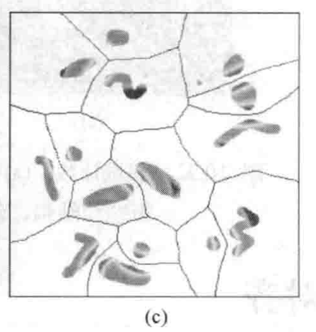


图 10.20　确定最小间距：(a)显示有几个特征的图形；(b)泰森多边形曲面中，特征的 Skiz 骨网和各单元的欧氏距离图；(c)赋给特征中各像素的 EDM 值（以伪彩色显示）。每个特征的最小值是该特征到其最近邻特征的边缘间距离的一半

测量最小间距的第二种方法是，构造背景的 EDM 并沿 Skiz 骨网上的点找到局部极小值。它们是

是最接近的点的正中间，提供了特征之间最小间距的另一种测度。使用这种技术，可分别测量一个特征上的不同点与一个或多个最近邻特征上的不同点之间的多个最小间距。

在空间填充结构（如组织中的细胞、金属中的颗粒或航测中的土地）情形下，人们感兴趣的是相邻特征的数量。由于这些特征是分离的，因此必须有一条背景像素线来分隔它们。在许多情形下，产生这条线的方法是，对经阈值处理的边界图像骨加化，然后将结果反转来定义像素。统计每个特征的相邻特征的数量的方法是，检查与边界中每个背景像素相接的那些像素的特征识别号。构建彼此相邻特征的表格，以便计算此类相邻特征的数量。

根据邻域数来标记特征，可揭示出结构的一些有趣性质。图 10.21 中标记了硅藻中的邻域数量，硅藻实质上是一种二维结构，其中的多数细胞形状为六边形。在彩色编码后的图像中 5 邻域和 7 邻域细胞配对很明显，但在原始图像中看不到这种配对。图 10.22 显示了对三维金属结构中微粒的一种类似标记，就如同在二维截面上揭示的那样。对于完全再结晶金属中的均衡结构，最近邻分布应是对数正态的（就如例中所示的那样）。

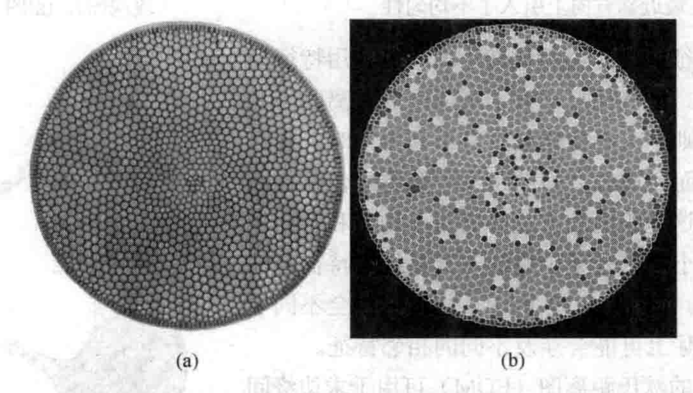


图 10.21　邻域计数：(a)硅藻图像；(b)根据边数着色后，结构中的孔洞。主要形状为六边形，但有些为五边形和七边形，它们通常彼此相邻

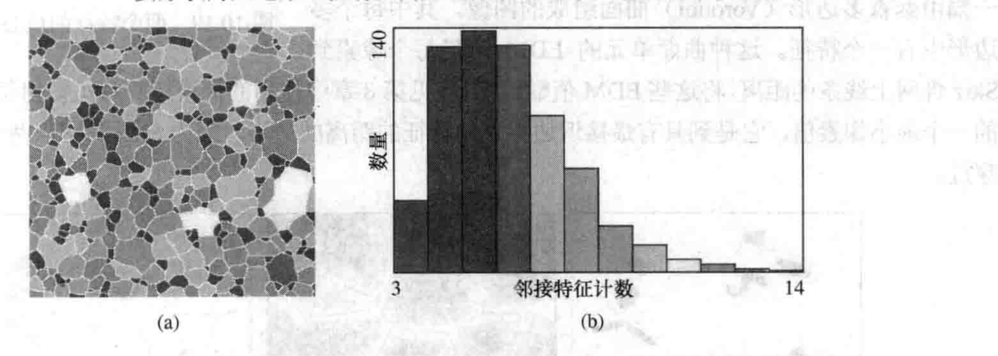


图 10.22　邻域计数：(a)已根据最近邻域数量着色的钢合金中的颗粒；(b)对于平衡颗粒结构，邻域数量的频率分布在二维和三维情形下是对数正态的

10.5　对齐

人们擅长于目视对齐和排列图像中的特征，但有时会发现这样的对齐和排列并不真正存在。第 2 章以太空中的星座为例，说明了对无序强加秩序的做法。

在许多图像分析情形下，需要有一个算法程序来确定对齐和排列。最常见的一种情形是接上断线，或将一系列点拟合为一条线，特别是直线。在所有分辨率下，这样的程序非常有用，如在侦察照片中

定位输电线路（照片中显示了塔架，但不能分辨出电线），在透射电子显微镜中勾画出原子晶格。通常，沿线出现的点的间隔并不规则，且可能未准确位于线上。这种不规则性和对视野中以其他点或特征形式出现的噪声的敏感性，就带来了问题。

将图像变换到不同空间中可解决此问题。在第 6 章给出了傅里叶空间的用法，但它对此目的并不适用，因为点的间距并不规则。好在还有其他我们不太熟悉的变换空间。将图像数据变换到霍夫空间可发现对齐（Hough，1962；Duda & Hart，1972；Ballard，1981）。不同的霍夫空间可用于匹配不同类型的形状，且在多数情形下，需要具备关于线型或其他排列的先验知识，以便对数据进行拟合。下例显示了一条直线的情形，因为这是最简单的情形和广泛使用的情形之一。

在图形上将一条线拟合到数据点的常规方法是“最小二乘”法，即使得每个点到线的偏差的平方和最小。霍夫方法可自动完成这一任务，因为它最小化了各个点到线的偏差，同时能正确地处理点不沿线均匀分布的情形。

定义一条直线需要两个参数。在笛卡儿坐标系（直角坐标系）中，直线方程为

$$y = mx + b \tag{10.9}$$

式中，m 是斜率，b 是截距。因为对于平行于 y 轴的直线，m 会变得无穷大，因此在霍夫变换中通常并不采用这一表示，而采用直线的极坐标表示。极径 ρ 和角度 ϕ（从原点到直线的法线长度和角度）定义了该直线。角度可在 $0\sim2\pi$ 间变化并保持 ρ 为正，或极径 ρ 可正可负而角度 ϕ 在 $0\sim\pi$ 间变化。下面的例子中将使用后一种约定。图 10.23(a)显示了实（或像素）空间中的一个黑点，转换后生成了图 10.23(b)所示霍夫空间中的正弦曲线。沿正弦曲线出现的每个点对应于穿过原始点的单条线的 ρ、ϕ 值。图中以彩色显示了这样的几个点，霍夫空间中的彩色点对应于实空间中的相应直线。

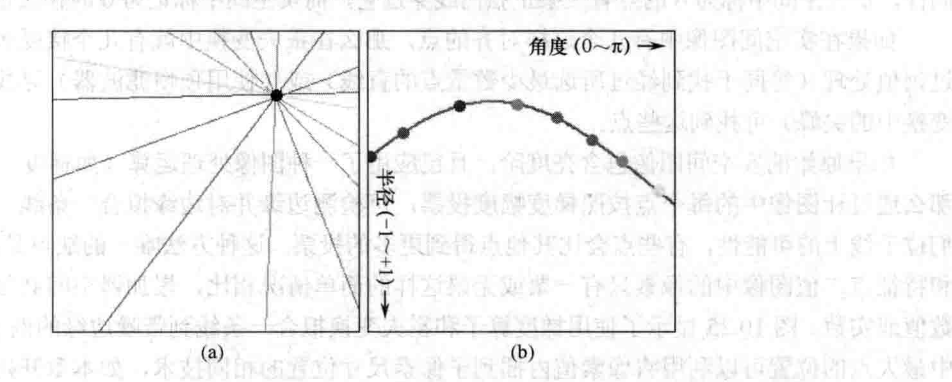


图 10.23　霍夫变换的原理：(a)实空间图像中的每个点生成霍夫空间中的一条正弦曲线；(b)实空间图像中的每个点表示可通过它画出的所有直线。霍夫空间中的每个点对应于实空间中的一条线，实空间的线对应于沿正弦曲线分布的一些点，为使这些线与点匹配，已对它们着色

霍夫空间是一个累加器空间。这意味着它会对图像中的许多像素的投票数求和，以便使霍夫空间中有较大总投票数的点可解释为实空间图像中的相应对齐。对于直线拟合数据的线性霍夫变换，该空间是具有角度和极径坐标的一个数组单元（像素，因为这是另一幅图像）。霍夫空间中的每个点定义了对应于角度 ϕ 和半径 ρ 的一条可在实空间图像中画出的直线。尽管这里是针对对齐来说明霍夫空间及其构建和使用的，但正弦（极坐标）形式的霍夫空间与断层摄影中所用的雷登变换相同，详见第 13 章。

为了构造霍夫变换，实空间图像中出现的每个点要将其票投给霍夫空间中穿过该点的一条线。如图 10.23 所示，这意味着实空间图像中的每个点在霍夫空间中生成了一条正弦曲线。沿霍夫空间中正弦曲线分布的每个点得到一票，并加到实空间图像中的每个点上，或根据与实空间中该点相关的亮度进行小数投票。在实空间中叠加来自几个点的正弦曲线，会使得票数加到它们相交的位置。霍夫空间中的这些交点出现在识别穿过实空间图像中多个点的直线的 ρ 和 ϕ 值处，如图 10.24 所示。

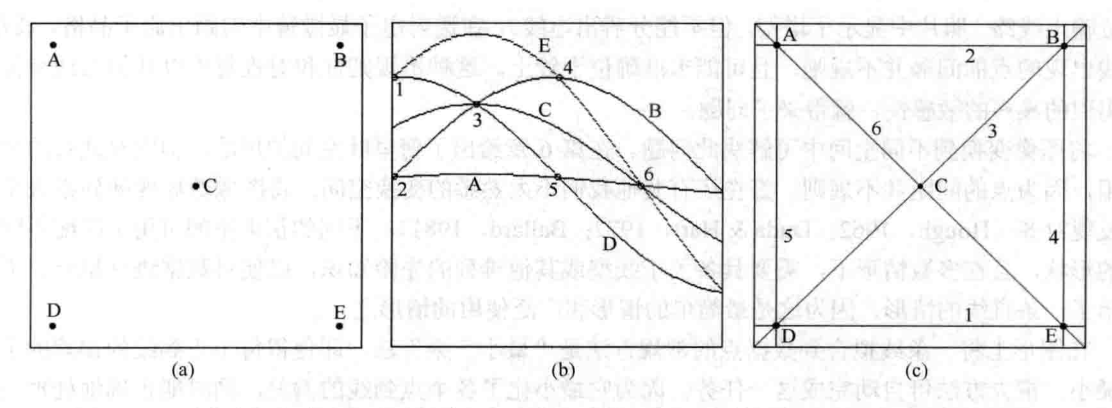


图 10.24　霍夫空间和实空间中线和点的对偶性。实空间图(a)中的每个标记点在霍夫空间图(b)中产生一条正弦线。对该空间中的交点编号，会生成实空间图像中的几条经过原始点的直线(c)。标记 A～E 识别实空间图像中的点（及它们在霍夫空间中的对应正弦线），而标记 1～6 识别霍夫空间图像中的交点（及它们在实空间中的对应直线）

在该例中，强调了实空间和霍夫空间中线和点的对偶性。实空间中 5 个原始点（标为 A～E）中的每一个，在霍夫空间产生一条正弦线（类似地进行了标记）。在这些正弦曲线的相交处，它们识别出霍夫空间中的点（标为 1～6）。这些点对应于实空间中经过同样这些点的直线（类似地进行了标记）。例如，在实空间中有三条线穿过点 A，而霍夫空间中标为 A 的正弦曲线会经过这三条线中的每一条的这些点。同样，霍夫空间中标为 6 的点有三条正弦曲线穿过它，而实空间中标记为 6 的相应直线穿过三个点。

如果在实空间图像中有几个这样对齐的点，那么在霍夫变换中就有几个接受多个投票的位置。通过阈值处理（等同于找到经过所选最少数量点的直线）或（使用顶帽滤波器）寻找局部极大值（霍夫变换中的尖峰）可找到这些点。

如果原始的实空间图像包含亮度阶，且已应用了一种图像处理运算（如梯度，即 Sobel 滤波器），那么通过让图像中的每个点按照梯度幅度投票，可检测边缘并对边缘拟合一条线。这意味着，根据它们位于线上的可能性，有些点会比其他点得到更多的投票。这种方法唯一的缺点是，使用这类投票时，和特征点二值图像中的像素只有一票或无票这样的简单情况相比，累加器空间必须能够处理更大的整数值或实数。图 10.25 显示了使用梯度算子和霍夫变换拟合一条线到带噪边缘的例子。注意，霍夫空间中最大点的位置可以利用将像素值内插到子像素尺寸位置的相同技术，如本章开始所述。也可使霍夫空间图像如所需要的那样，大到足以提升 ρ、ϕ值的精度。

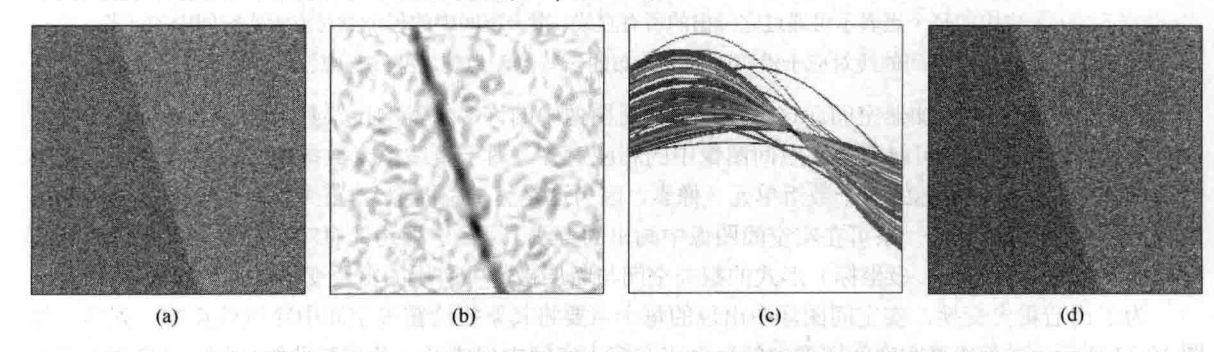


图 10.25　使用霍夫变换将一条直线拟合到噪声边缘：(a)原始带噪图像；(b)平滑图像降噪并应用一个 Sobel 梯度滤波器后的结果；(c)梯度图像的霍夫变换；(d)霍夫变换中最大点定义的直线

图 10.26 显示了该技术的一个应用。任务是求出图 10.26(c)所示监控视频图像中的人的身高。该

计算的一个方便标尺是桌子的已知高度。图 10.26(a)显示了视频中的一幅较早图像，其中可以看见桌子。对边缘周围的区域应用 Sobel 梯度滤波器，突出了绿色标记的像素。使用这些像素生成霍夫变换［见图 10.26(b)］来拟合红色标记的线。将这条线叠加到人像上，就可测量该位置桌子的高度，然后结合桌高与人高（并对相机视角进行一个小的校正），就可进行高度测量。

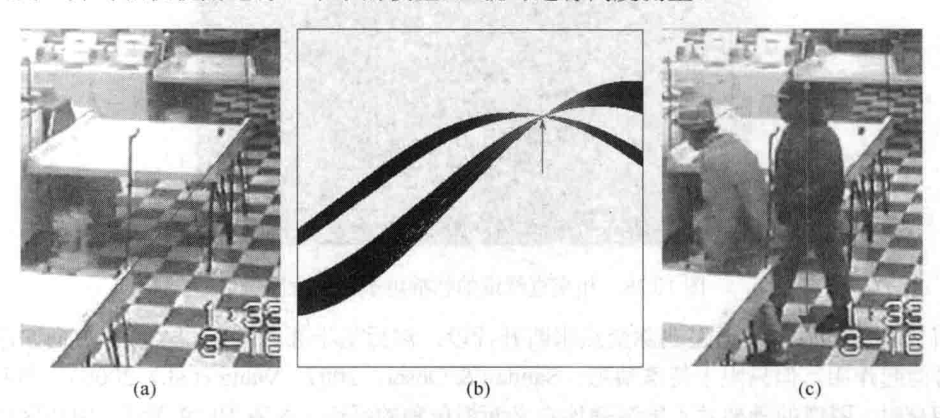


图 10.26　监控视频图像中的高度测量：(a)参考图像，其中的桌子用做校准标尺。Sobel 滤波器突出了绿色像素，然后使用这些像素生成霍夫变换。结果是红色直线；(b)霍夫变换，显示了由 Sobel 边缘滤波器突出的像素所生成的正弦曲线，以及这些曲线的交点，产生了一个高累积投票（红色箭头）；(c)叠加到待测人体身高图像上桌子边缘间的直线。由绿色和蓝色箭头长度之比，可测出身高。需要对相机到蓝色箭头中心的较大距离（桌高）与相机到绿色箭头中心的距离（人高）之比进行透视调整

对于平行线间距的测量，霍夫空间将问题简化为测量变换中的尖峰间的垂直距离。如图 10.27 所示，对集成电路上一条线的 SEM 图像求导，在同一ϕ值（对应线的方向）产生尖峰，测量ρ上的距离可直接测量线宽和线间距。类似地，线的角方向测量是通过确定变换中水平轴上的ϕ值来完成的。当有许多点需要拟合时，其中的一些点要比其他点重要，因此按比例投票特别有用，就如使用梯度算子来定位边缘时那样。按照亮度（或由亮度得到的某些值）对投票数求和，可让某些点的权重更大，进而更好地拟合结果。

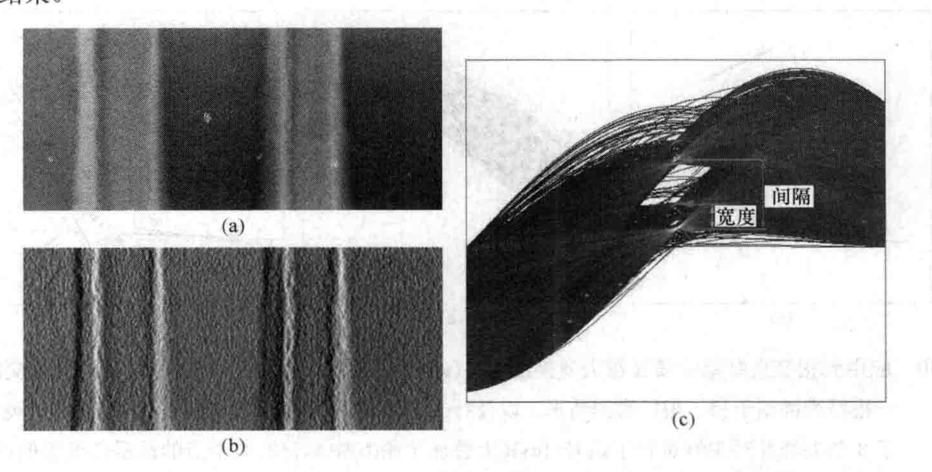


图 10.27　测量集成电路上的线宽和线距：(a)原始 SEM 图像；(b)水平求导后的结果；(c)霍夫变换，显示了对应于线边缘的尖峰和测量宽度与间距的距离

线性霍夫变换还可求解图像分析中经常出现的一个问题，如图 10.28 所示。在各种工业场合制备直纤维标本时，往往会得到纤维彼此相交或彼此靠近的结果。

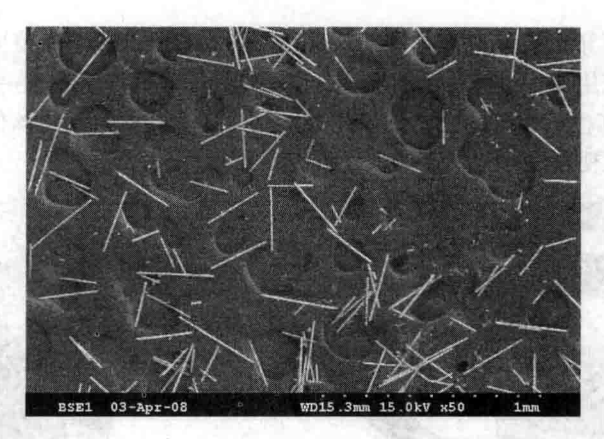


图 10.28　相交直纤维的扫描电子显微镜图像

使用纤维图像的骨架，通过删除交点来断开片段，然后基于邻近端点和类似的方向来匹配各片的策略，有时能起作用，但只限于特殊情形（Sandau & Ohser，2007；Wang et al.，2008）。当两根纤维以较小角度相交时，图像的骨架并不能清楚地定义两根单独的纤维，如图 10.29 所示。中心区域两根纤维重叠的骨架，与两根纤维的方向均不能精确匹配，但必须包含在二者的测量中。

处理此类图像的一种不同方法是使用线性霍夫变换。如图 10.30 所示，纤维图像的变换中包含了所有纤维中每个像素的正弦曲线，它们的总和会产生极大点，这些点对应于每根纤维的方向和位置。此外，这些极大点的值表示了对这条线有贡献的像素数量，因此可测量各根纤维的尺寸。如果需要，沿由变换确定的每条线扫描，可以定位每根纤维的端点。

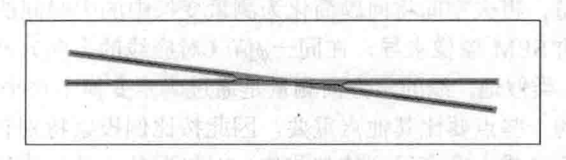


图 10.29　两根相交纤维图示。在两根纤维的测量中，必须包含中心区域的骨架线段

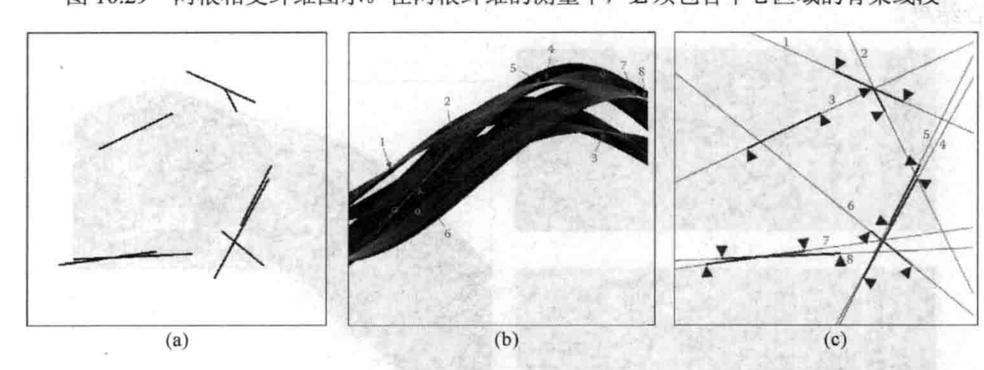


图 10.30　应用到相交直纤维的线性霍夫变换图示：(a)8 根纤维，包含了以较大和较小角度相交的情形，一根纤维终结于另一根纤维的情形，以及纤维近距离平行的情形；(b)图(a)的线性霍夫变换，标记了 8 个尖峰并对尖峰进行了编号；(c)其上叠加了图(b)中标记的 8 个点的反霍夫变换的原始图像，其中每个编号的最大点有一条对应直线。蓝色箭头表明每根纤维的端点由对直线的扫描确定

对其他形状可直接应用霍夫变换，但用来累加投票数的霍夫空间的大小和规模会随形状的复杂性而增大。使用一个圆来拟合一组点，可测量电子衍射图样并定位机器零件上的钻孔（且原理上可在内布拉斯加州的卫星图像中定位并测量中心旋转灌溉系统）。并不要求这些点形成一个完整的圆；例如，该方法一直被用于测量陶罐碎片的曲率。

圆形霍夫变换需要一个三维空间，因为定义一个圆需要 3 个参数（圆心的 x 坐标和 y 坐标以及半径）。如图 10.31 所示，实空间图像中的每个点在霍夫空间中产生一个圆锥形投票，它对应于可通过该点画出的所有圆，这些圆的半径和圆心位置可变（Bresenham，1977；Kimme et al.，1975；Rizon et al.，2005）。

如果只知道其中的一个或两个值（如电子衍射图样的中心或钻孔的半径），或知道这些值的值域很小，则可降低霍夫空间的规模和大小，使得整个过程更为高效。准确地将一个圆拟合到一组不规则间隔且亮度变化的点，是霍夫方法测量图像能力的一个很好示例。由于从定义该圆的霍夫空间中的多个投票生成的点的亮度，也是从实空间中对应圆上的所有点的那些投票之和，所以这种方法也可有效地整合电子衍射图样中各点的总亮度。

图 10.32 显示了采用圆形霍夫变换来在选定区域衍射图样内定位“最佳拟合”圆的一个例子。图样本身有围绕在每个圆周围不规则位置的亮斑。人不难估计圆的位置，但通常很难让计算机算法来检查局部像素区域。霍夫变换可找到圆的位置，且可很容易地测量围绕该圆的整体亮度。过图案中心且平行于极轴的直线上的值，可给出图案的整体径向亮度图。人们一直使用这种方法来从识别具有很多斑点的电子衍射图样中的石棉纤维（Russ et al.，1989）。

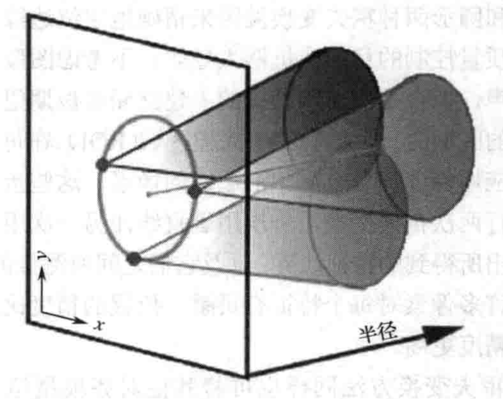


图 10.31 圆形霍夫空间。图像平面上的 3 个点（蓝色）都可生成一个圆锥体。圆锥体的交点（红色）指定半径并定位通过原点的圆心（绿色）

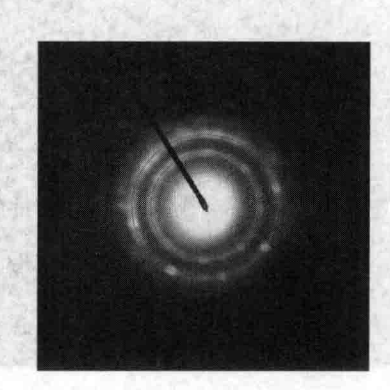
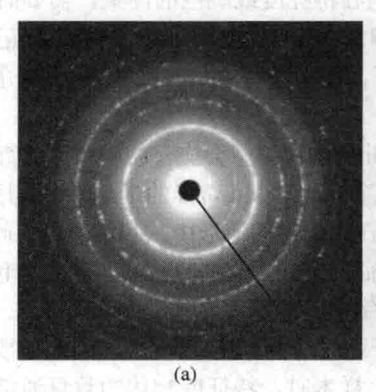
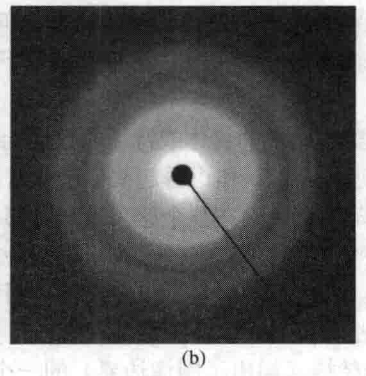


图 10.32 选取的电子衍射图样区域，其中主圆（及其圆心）可由圆形霍夫变换定位

测量衍射图样（以及类似的任务）中的极径和点的整体亮度的另一种方法是，画出像素总亮度与极径的关系图。此时需要找到圆心的位置，而找到圆心的位置则需要使用圆形霍夫变换。如图 10.33 所示，减去应用形态学灰度开运算删除各个斑点后的背景，只留下了斑点。绘制像素平均亮度图需要对亮度值求和，并将其计数为极径的函数。



(a)



(b)

图 10.33 径向亮度剖面：(a)原始电子衍射图样；(b)灰度形态学开运算后得到的背景；(c)图(c)减去图(b)后的结果；(d)径向亮度剖面

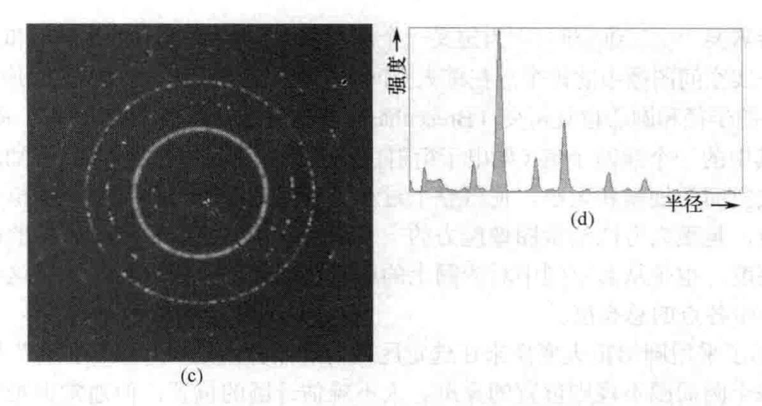


图 10.33　径向亮度剖面：(a)原始电子衍射图样；(b)灰度形态学开运算后得到的背景；(c)图(c)减去图(b)后的结果；(d)径向亮度剖面（续）

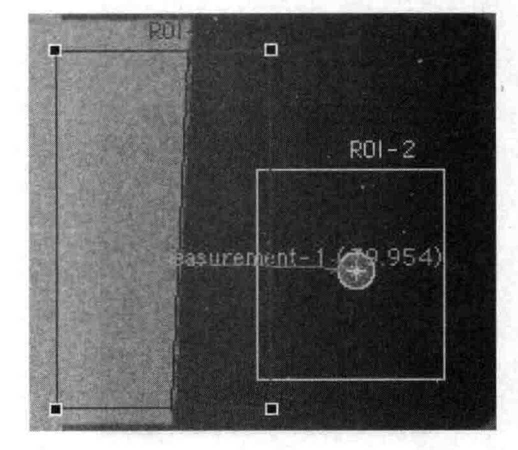


图 10.34　使用线性和圆形霍夫变换来对加工零件的边缘和圆孔进行定位，进而以亚像素精度测量它们间的距离

图 10.34 显示了来自机器视觉的一个例子，其中的线性和圆形两种霍夫变换被用来精确地定位边缘和洞以用于质量控制的目的或机器人导航。不考虑图像中的总体噪声，拟合是快速和鲁棒的。建立涵盖被期望的每个特征的区域的一个单独的感兴趣区域(ROI)。在每个 ROI 中，应用梯度算子来辨识沿边缘的像素。这些接下来用来进行两次霍夫变换，一次用于直线而另一次用于圆。显示出所得到的特征边界，以及它们之间所测量的距离。由于许多像素对每个特征有贡献，位置的精度比像素规格的精度更高。

霍夫变换方法同样也可将其他对齐模型拟合到数据。但约束条件是，随着代数形式的模型变化，每个可调的常数要求霍夫空间中的另一个维度。构造变换并找到最大点需要更多的内存。

10.6　计数

在许多图像中，统计物体的数量要求识别某些独特亮度或颜色的图案，需要时这些图案应可被计算机程序或人来识别。人通常不擅长精确计数，但却善于识别。因此，将图像标记为某些颜色并让计算机来计数通常是较为可取的策略。在某些情形下，使用某种方法（如第 6 章中介绍过的互相关）可便于自动计数。

统计图像或视野中出现的物体或特征的数量（此处的讨论中会交替使用物体或特征这两个术语），是最常用的图像分析过程之一。概念看起来似乎非常简单，但令人奇怪的是，许多程序却把它理解错了。错误与视野的有限范围有关。在整个感兴趣视野都在图像内的情形中，基本不会有问题。每个物体或特征都由一些点和这些点的数量定义。取决于识别特征的方式，有些算法会忽略完全包含在其他特征内的孔洞中的那些特征。在使用边界表示来定义二值图像中的特征边界时，最常出现这一问题。统计其他特征内部的特征数量是否合适，以及图像是投影还是截面，与特定应用有关，因此必须由用户来决定。

当视野是整个结构（超出了图像边界）的一个样本时，特征或物体的数量通常表示为单位面积上的数量。当特征与视野的边缘相交时，统计所有可见特征是不正确的。得到无偏结果的常用方法是，只统计接触两个边缘（如顶部和左侧）的那些特征，而忽略接触另外两个边缘（如右侧和底部）的那

些特征。这等同于按特征的右下角来统计特征的数量。由于每个特征有且只有一个右下角，因此统计这些点的数量等同于统计特征的数量。图 10.35 显示了一个例子。有时使用的另一种方法是，将与任何边缘相交的特征计数为 1/2。对于凸状特征，这两种方法对是等效的，但对于非凸状特征，由于会它们会与边缘产生多个交点，因此会出现问题。

“右下角”方法类似于通过统计鼻子的数量来确定房间内的人数。由于每个人只有一个鼻子，因此鼻子的数量与人数相同。如果在房间内只标记了一个较小的区域，那些人可能恰好位于边界上，此时统计鼻子的数量来确定人数仍然有效。鼻子在该区域内的任何人都被计数，而不管这个人的身体有多少在区域之外。相反，鼻子在区域之外的任何人都不计算在内，而不管这个人的身体有多少部分在该区域内。要特别注意的是，在该例中，在区域之外可以看到部分人体，但对于图像中的物体却并非如此。由图像获取的视野之外的物体的任何部分，按定义是看不见的，因此也就没有视野之外物体的数量和形状信息。这就是为将用于计数的每个物体定义一个独特点的原因。

并非所有系统都可统计只接触两个边缘的特征的数量。有些软件包提供了两个选项，即要么统计所有特征的数量而不管它们是否接触边缘，要么只统计未接触碰任何边缘的那些特征的数量。这两个选项都不正确。与统计物体的数量相反，测量物体需要一个更复杂的过程。与任何边缘都相交的特征是不能测量的，因为它未被全部成像，因此不能正确地获得其大小、形状或位置信息。若只统计图 10.36 中未与任何边缘接触的特征，那么大特征和小特征的比例就错了。大物体很容易接触边缘，因此不能测量与视野边缘相交的大特征的不成比例的部分。

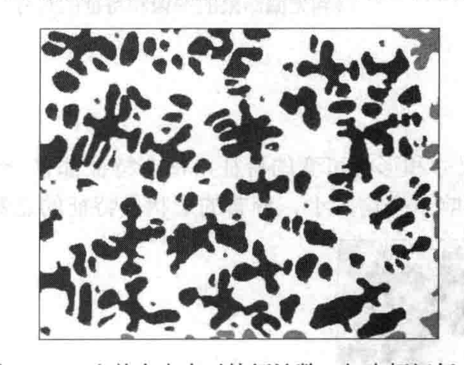


图 10.35　由单个点来对特征计数。红点标记每个特征中最低像素线的右端。统计接触顶部和左侧边界的特征，但不统计接触底部和右侧边缘的那些特征

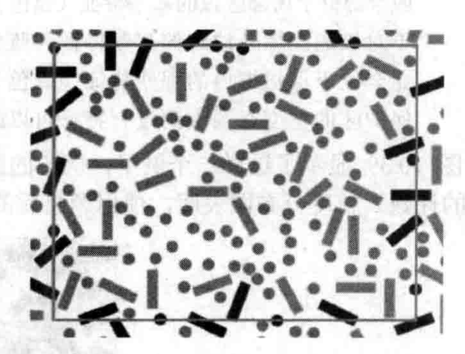


图 10.36　当一个图像区域（以绿框显示）位于待测物体的一个较大区域内时，有些物体会与边缘相接。因为这些特征伸出图像区域外部的部分是未知的，因此无法测量。在所示的例子中，尽管小的红色特征数几乎是长蓝色特征数的 3 倍，但与边框相交的特征数量却是蓝色特征数量的 2 倍，因此结果出现了偏差

校正这种编差的方法有两种。两种方法得到的结果相同，但实现方式不同。最初用于手工测量相片的较老方法，是在图像内设置一个“保护框”，如图 10.37 所示。和前面介绍的一样，这种方法并不统计或测量视野中接触下部和右侧边缘的那些特征，只统计或测量跨越保护框顶部和左侧边缘的那些特征。不统计只处在保护区域内的特征，无论它们是否接触视野的实际边缘。所统计的特征或物体数量就是单位面积上的无偏测量结果，但该面积是计数框内的面积，而不是图像的整个面积。这完全等同于上面介绍的“右下角”计数方法，但也允许对物体进行测量。由于保护区域必须要宽到足以保证没有特征伸到保护区外，因此与整个图像区域相比，该活动区域可能会明显减小。

第二种方法使用整个图像区域并测量与所有边缘都不接触的那些特征。为补偿因较大特征更可能接触边缘而在测量过程中将其忽略所导致的偏差，应按照特征可能接触随机放置在视野内的边缘的可能性，按比例来统计这些特征的数量。如图 10.38 所示，调整后的每个特征计数为

$$\text{计数} = \frac{W_x W_y}{(W_x - F_x)(W_y - F_y)} \tag{10.10}$$

式中，W_x 和 W_y 是图像在 x 和 y 方向的尺寸，F_x 和 F_y 是特征在这两个方向上的最大投影尺寸。

F 值是前面介绍的外接矩形坐标。当特征尺寸比视野尺寸小时，该分数接近于 1.0，计数不受影响。当特征在任何一个方向伸出视野时，在样本上随意放置该视野更可能使得特征与边缘相交；因此，必须统计那些可被测量的特征，因为对这些特征的多次校正可能已使得它们消失不见。调整的计数因子可进行这种补偿。

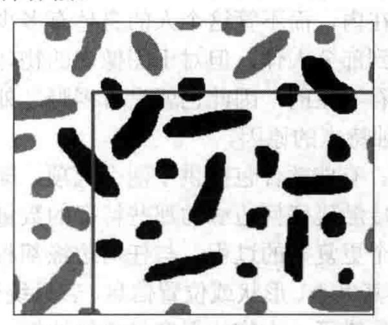


图 10.37　使用保护框（显示为红色）进行测量。对部分或全部位于内部区域的那些特征（黑色）计数。不对接触图像外边缘的那些特征（蓝色）或完全位于保护框内的那些特征（绿色）计数。保护区的宽度必须大到没有特征可以跨越它

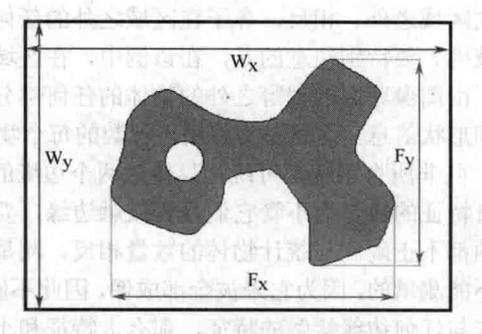
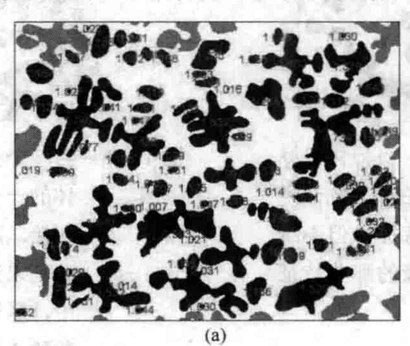
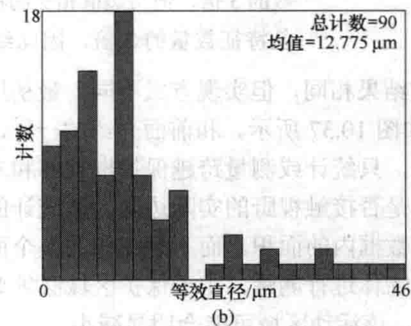


图 10.38　在式（10.10）中用于调整计数来得到无偏结果的图像和特征的尺寸

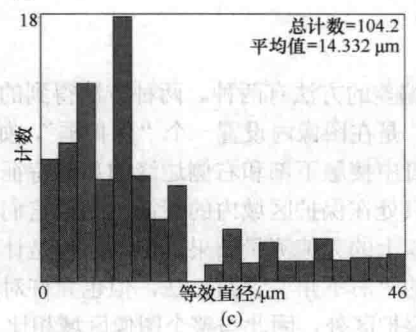
图 10.39 显示了这样一个例子，例中的图像包含了尺寸和形状可变的特征，每个特征都有一个调整后的计数。粒度分布图表明，使用调整后的计数明显影响了平均大小、分布的形状及特征的总数。



(a)



(b)



(c)

图 10.39　对不同大小特征调整计数的例子：(a)每个特征调整后的计数；(b)未校正因特征与边缘相交导致的偏差时的粒度分布数据；(c)使用经调整计数后的结果。注意特征总数（90μm 对 104.2μm）与平均特征尺寸（12.8μm 对 14.3μm）间的差别

10.7 特殊的计数过程

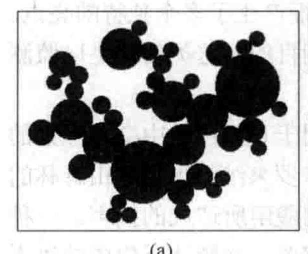
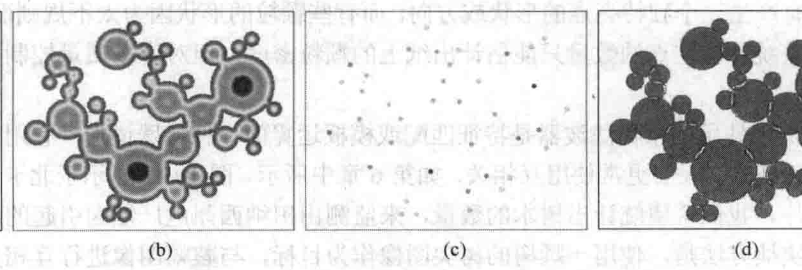
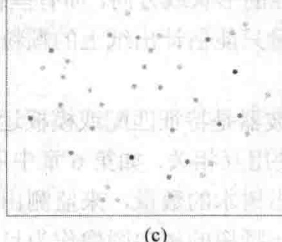
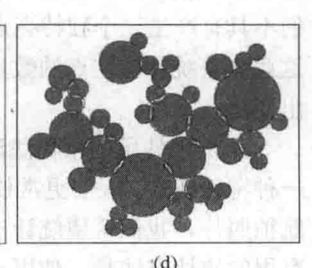
上一节的计数示例中，默认假设了特征是分离且不同的。在前几章中，给出了处理彩色图像、灰度图像和经阈值处理后的二值图像的处理过程，目的是分离出特征以便进行计数和测量（如分水岭分割）。但也存在使用这些方法不能成功或很难进行计数与测量的情形。若有足够多关于标本的独立知识，那么在这些困难情形下也可进行计数。

下面以第 8 章中的图 8.70 所示的交叉纤维为例进行说明。这类结构在生物样品、造纸所用的木质纤维、纺织、食品加工等情形下很常见。由于纤维相交，从图像分析算法的观点来看，只出现了单个特征，且该特征通常会在视野的所有方向延伸。但我们可估计出单位面积上的纤维数量和平均纤维长度。通过对纤维进行阈值处理并对得到的二值图像骨架化，可显示一系列相交的中线。如第 8 章所示，纤维的端点可以统计为骨架上恰好只有一个相接邻域的那些像素。这种方法并不完美，因为真实纤维的某些端点会被隐藏，这不是一种完美的方法，有些纤维的端点会因为其正好位于其他纤维上方而被隐藏，但原理上可以检测到这些端点，因为这些端点会生成恰好有 3 个邻域的像素，且相交的角度约为 180°。但即使没有这种细化，端点数的一半也可给出近似的纤维数量。

在上一节中，每个特征是由一个独特的点来计数的。计算纤维的端点数时，为每根纤维使用了两个独特的点。如果纤维的一个端点在视野之内，而另一个端点在视野之外，则将其计数为 1/2 根纤维，因为另一个端点会在不同的视野中计数。将计数除以图像面积，可得出单位面积上的纤维数量。纤维总长（骨架化后中线的总长）除以纤维数量（端点数的一半），可得出平均纤维长度。如果纤维长度远大于视野外范围，以致许多视野中均不出现端点或很少出现端点，就需要组合来自多个视野的数据来获得统计上有意义的结果。

第 8 章中介绍了分开相接特征的几种技术。最好的分水岭分割技术利用了欧氏距离图。该运算对特征内的每个像素按该像素到最近背景点的距离大小，赋一个灰度值。然后使用距离图中位于两个最高像素之间的谷或点，来定位可在特征之间画出的边界，进而分开这些特征。因为这种方法假设围绕特征聚类边缘的主要凹口是一个分离点，因此该技术也称为凸分割。

如果处理和分割的目的是对特征计数，那么有一种可节省大量计算的快捷方法。在欧氏距离图（EDM）中，每个局部极大点（大于等于其所有 8 个相邻像素的那些像素）是一个独特的点，它表示可通过分水岭分割最终与相接特征分开的一个特征。找到这些最终的腐蚀点（UEP）可快速统计出所出现的特征的数量。此外，EDM 中该点的灰度值是相应特征大小的一个测度，因为它对应于可在该特征内画出的最大内切圆的半径。图 10.40 再次显示了第 8 章中的一些相接圆的例子，其中给出了从欧氏距离图中找到了最终的腐蚀点和相应圆的大小。

(a) (b) (c) (d)

图 10.40 统计相接颗粒的数量并测量其大小：(a)相接圆的二值图像；(b)欧氏距离图（色彩表示了像素到最近背景点的距离）；(c)最终的腐蚀点（欧氏距离图中的局部极大），其颜色给出了半径；(d)基于最终点的圆形颗粒边界相交情况

图像分析算法在统计相接颗粒数量时会遇到许多困难。如果这些颗粒的形状不为凸状，以及颗粒不在单个平面上且只有边界相接，那么分水岭分割就可能无法成功地分开它们。实际需求是，特征的形状产生了只有一个尖峰的 EDM，这意味着有些特征（如图 10.41 中的那些不完全呈凸状的特征）可以成功地分离，但情形并非总是如此。

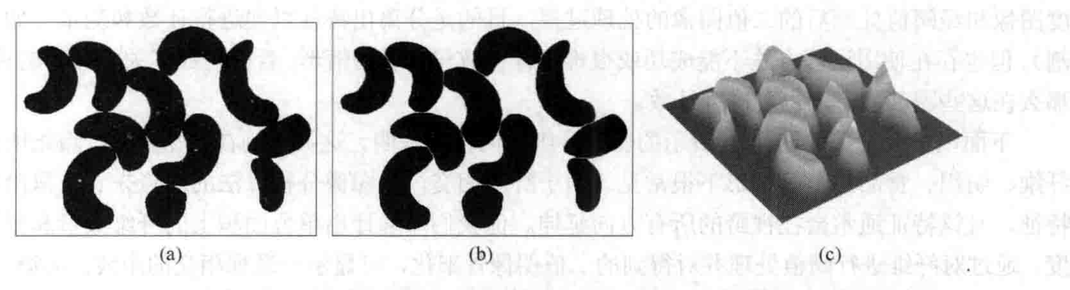


图 10.41　非凸状特征的分水岭分割：(a)相接腰果经阈值处理后的图像；(b)成功进行分水岭分割后的图像；(c)欧氏距离图，它表明每个相接的特征都只有一个极大点

图 10.42 中显示了更难处理的一个典型例子。涂在纸张上的黏土颗粒层基本上只有一个颗粒厚，但这些颗粒大量重叠。为计算出纸张单位面积上的颗粒数，需要再次使用为每个颗粒找到一个独特点的想法。

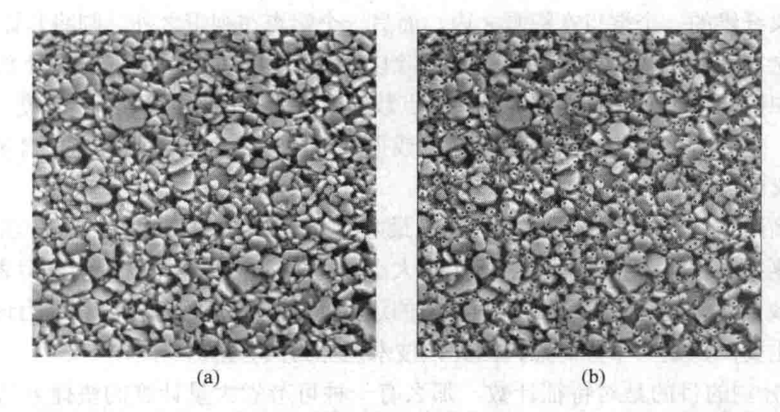


图 10.42　统计重叠颗粒的数量：(a)纸上黏土颗粒的 SEM 图像；(b)使用顶帽滤波器找到局部最亮点并将它们叠加到原始图像上的结果，使用这些点可计算出每平方英寸面积上的颗粒数量

此时，扫描电子显微镜形成图像的方式可提供一个答案。多数颗粒都有角和一个最高点，这个最高点在二次电子图像中显示为亮度。使用第 5 章中介绍过的顶帽滤波器可分开这些点（如图所示）。在该例中，顶帽滤波器为多数黏土颗粒找到了具有亮度局部极大的一个独特点。有些点未找到，因为它们不具有产生一个独特亮点的形状或方向，而有些颗粒的形状因为太不规则而产生了多个独特的亮点。这意味着统计这些点的数量只能估计出纸上的颗粒密度，但对许多质量控制目的，这类估计足以监测出变化。

上例中使用的顶帽滤波器是特征匹配或模板运算的一种特殊情形，它用于查找图像中亮暗像素的一种特殊模式。人们更常使用互相关，如第 6 章中所示。图 10.43 显示了北卡罗来纳州米切尔山森林的航拍照片。我们希望统计出树木的数量，来监测由田纳西州电厂烟囱引起的酸雨所造成的损害。一种有用的快捷方法是，使用一颗树的树尖图像作为目标，与整幅图像进行互相关，并统计所生成的斑点数量。当然，这种方法是近似的，它确实可以找到部分隐藏在另一棵树后面的树，或找到另一棵树前面的树，但会错过那些和目标例子具有不同树尖的树。

图 10.44 显示了另一个例子。颗粒的边缘不是亮度均匀或轮廓分明的，因为存在相邻的颗粒，有

些颗粒还会相接。此外，微粒大小差别很大，但每颗微粒中心附近的亮度变化是相当一致的。在一颗有代表性的颗粒中心选择一个区域（如图所示）作为与整个图像互相关的目标，产生了所示的结果（已应用伪彩色使得灰度亮度的细微变化更加明显）。对这幅图像进行阈值处理，并标记每个区域的质心，就识别出了用于计数的颗粒。但这幅图像的另一个问题是，被检查的表面不是平的且标本有深度，所以顶层下方有更多的颗料未被计数。尽管可自动地统计出标记颗粒的数量，但该数值的含义并不清晰。

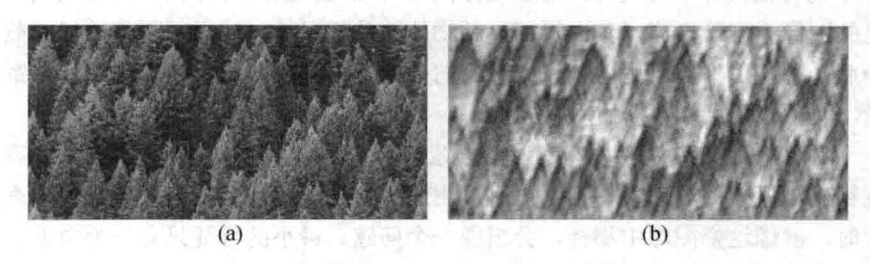


图 10.43　使用互相关计数：(a)树的航空照片，红框将所选的树顶标记为进行匹配的目标；(b)互相关结果，已标记出所识别的树顶

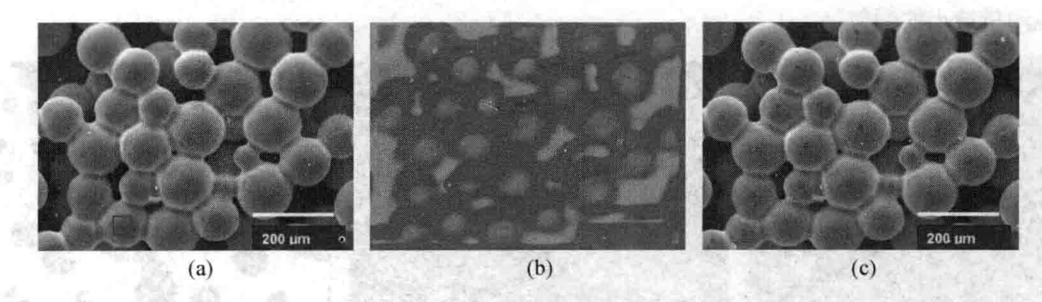


图 10.44　互相关计数：(a)颗粒的 SEM 图像，显示了作为目标的选取区域；(b)互相关结果，图像已伪彩色化来显示对应于颗粒的橙色孤岛；(c)橙色区域经阈值处理后的质心，已叠加到原始图像上

三维问题通常必须由三维成像或体视学关系来求解。图 10.45 显示了一个可成功计算聚类中颗粒数量的方法。这是碳黑颗粒的一幅 TEM 图像，碳黑颗粒以大小不等的聚类形式出现。若假设所有聚类中颗粒的大小和密度与孤立颗粒的大小和密码类似，则可由聚类的整体密度估计出聚类中的颗粒数。测量孤立颗粒的整体光密度的平均值（或取几个此类测量结果的平均值），并用每个聚类的整体密度除以该值所得到的数字，就是所出现颗粒数量的一个估计。

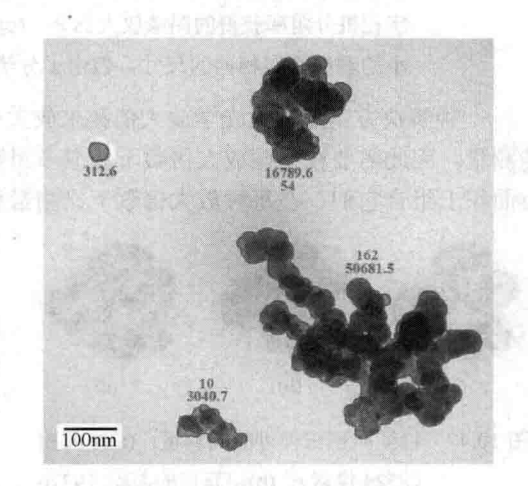


图 10.45　透射电子显微镜观察到的碳黑颗粒的聚类。将每个聚类的整体光密度（显示为红色）除以单个颗粒的整体密度，可估计出每个聚类中的颗粒数（显示为蓝色）

这种方法也可处理透射光图像和 X 射线图像。使用比尔定律（质量密度的指数吸收）设置图像的背景亮度值，使得零密度和最暗值对应于一个任意的正密度，可很容易地建立一条密度校准曲线。由于只使用了比率，因此不需要实际的校准标准。当然，要假设图像中没有点完全是黑的。

10.8 特征尺寸

图像中特征尺寸的最基本测度是面积。对于基于像素的表示，这是特征内像素的数量，它通过计数求出。对于边界表示，面积可由式（10.4）计算。对于游程编码的表示，它是特征弦长的总和。

二维图像中的特征尺寸以各种方式与三维空间中对应物体的尺寸相关，具体取决于图像是如何获得的。最常见的图像类型是投影，其中特征显示物体的外部尺寸，或显示物体的平面截面，即穿过物体的切片。在后一种情形下，可用前一章中讨论的体视学方法来估计物体的体积。本章介绍图像中特征尺寸的测量，而不介绍尺寸数据的后续分析。

图 10.46 显示了在平坦基底上散布有球形颗粒的一幅图像。在每个特征中必须要有足够的像素才能得出该特征精确测量值的约束条件下，可直接由这样的一幅图像测量出球形颗粒的直径。当颗粒的覆盖范围较大时，就像这幅图像中那样，会出现一个问题。最小的特征只有一个像素大小或几个像素大小，因此轮廓不分明（事实上有些特征甚至不能进行阈值处理）。需要一台高分辨率相机或组合多幅图像来在同一时间对大特征和小特征成像。此时，即使是高分辨率数码相机也不能提供足够的像素来准确地测量最小的颗粒。

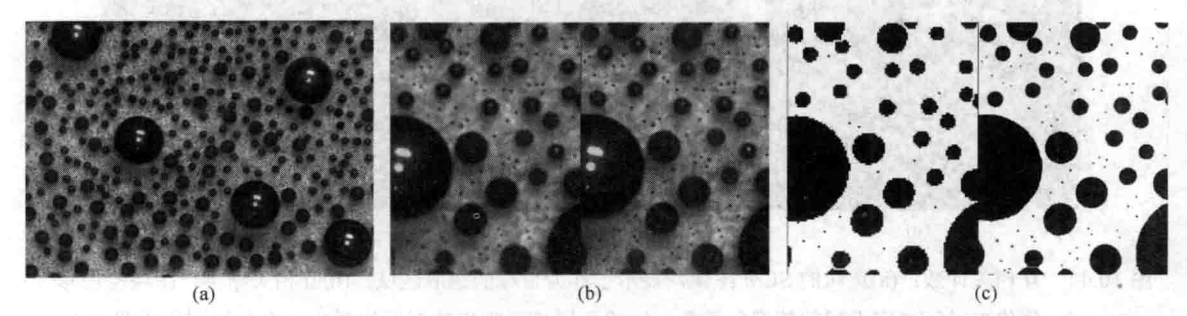


图 10.46　分辨率对特征测量的影响：(a)散布在平坦表面上的球形颗粒；(b)以视频相机分辨率和兆像素数字相机分辨率获得的图像放大部分；(c)经阈值处理、孔洞填充和分水岭分割后的一些区域。最小的颗粒都有相同的尺寸，但由于分辨率有限，该图像表示会变化，且不能充分地表示它们

一种解决方案是增大光学放大倍数来放大小颗粒，但此时大颗粒可能会因为与屏幕边缘相交而不能测量。因此需要在不同放大倍数下获取多组数据。以不同放大倍数进行测量时，正确的过程是在单位面积上组合它们。在每种放大倍数下只测量相同数量的图片会产生不正确的结果，因为在较大的放大倍数下，所覆盖的面积会小很多。

(a)　(b)　(c)

图 10.47　特征面积三种可能的测量：(a)净面积（8529 像素）；(b)填充后的面积（9376 像素）；(c)凸状面积（11227 像素）

即使是对统计像素来求出特征面积这样简单的思想，也需要做出某些决定。例如，考虑图 10.47 所示的特征。面积中是否应包含内部孔洞中的像素？当然，这取决于数据的预期用途。若孔洞是穿过物体内部空洞的一个截面，在特征面积与物体体积相关时，则应将其包含在内，而在面积与质量相关时，则不应包含在内。但我们很难知道孔洞是否是穿过物体上一个表面凹陷的截面，此时应将特征边缘凹陷处的那些像素也包含在面积内。如图中所示，这将产生三种不同的面积测量，即净面积、填充后的面积和凸状面积。

前两个面积的测量通常作为简单的像素计数练习来完成。在标记那些相互接触并包含特征的像素时，可检测出内部孔洞的存在并统计其内的像素。需要时，这些像素可加回到原始图像中来填充孔洞。

确定面积中是否包含这些像素就变成了基于其他知识的用户决策。

凸状面积的计算要稍复杂一些。在某些情形下，组合应用膨胀和腐蚀可为特征构造一个凸包并填充任何不规则的边界，进而通过像素计数来求出面积。但如第 8 章所示，采用方形像素网格时，这些方法会引起特征形状的失真。构造围绕特征的一个 n 条边的多边形的另一种方法，有时也称为特征的紧绷弦的或橡皮筋边界，因为它有效地定义了涵盖原始特征像素的凸状物的最小面积。

通过旋转像素坐标轴，可以找到任何方向上的最小点和最大点。旋转坐标系、对沿着特征边界的点计算新的 x', y' 值并搜索最小值和最大值的过程，是简单而有效的。对于任何特定的旋转角度 α，需要正弦值和余弦值。在多数情形下，这些值存储在一个表格中，该表格对应于程序中所用的特殊角度。然后可将新坐标计算为

$$x' = x\cos\alpha + y\sin\alpha$$
$$y' = y\sin\alpha - x\cos\alpha \tag{10.11}$$

以一系列旋转角度执行这一过程时，每个旋转坐标系中具有最大差异的点，就构成了上面所示的外接矩形的顶点。为构造凸状或紧绷弦轮廓，使用适度的旋转步骤可以得到有用的结果。例如，将轴线以 10° 为步长旋转，可得到一个有 36 个顶点和 36 条边的外接多边形。图 10.48 显示了这个过程，但只给出了几个步骤。多边形内的面积可以由式（10.4）用顶点坐标计算出来。

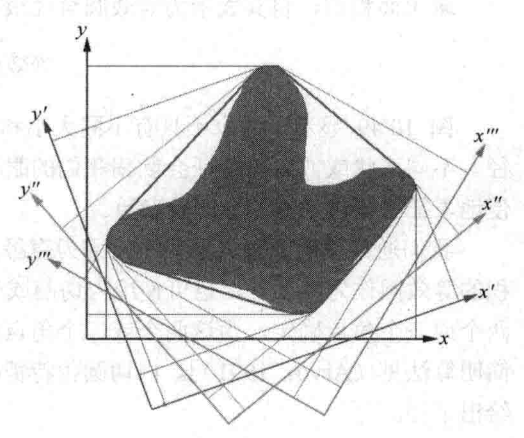


图 10.48 通过旋转 x、y 轴并定位最小点和最大点来构造外接多边形

图 10.49 比较了不同特征形状的外接多边形、等面积圆和最长轴。仅对于长而窄的特征，才需要更多的顶点来产生外接多边形，进而较好地近似外部部界。

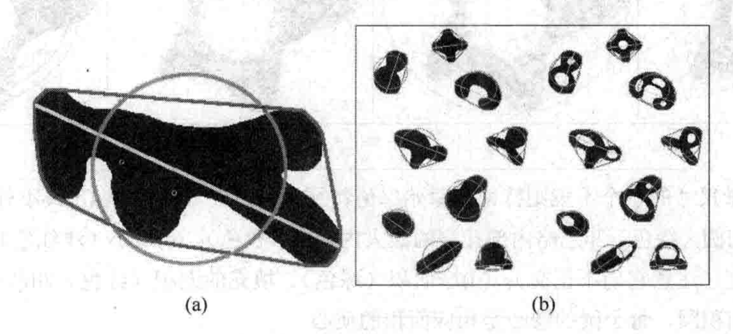


图 10.49 (a)不规则特征的外接多边形（红色）、最长弦（黄色）和等面积圆（绿色，圆心是特征的质心，圆面积与特征面积相同）的比较；(b)一种特征排列

但在定义和确定面积时，必须建立像素大小和真实结构大小间的转换因子。尺寸校准的方法是，获取已知标准特征的图像，然后使用稍后对未知特征应用的相同算法来测量它。对于宏观或微观图像，可用一个测量尺度。当然，要假设成像几何和光学不变；对于光学显微镜的玻璃镜片，这是一个好的假设，但不一定适合用于电子显微镜的电子透镜，也很难适合现实生活中的成像，此时必须对每幅图像进行合适的校准。例如，日常较好的做法是在犯罪现场照片中包含一把米尺。对于一些遥感运算，相机的位置和特性是已知的，且放大倍数是按几何方式算出的。如果放大比例在图像中因位置不同而不同，或因方向不同而不同，那么由于场景几何（如一个斜视角）或仪器的变化［如 AFM 扫描中的速率相关失真；见（Nederbracht et al.，2004）］，在获取有用测量结果时会出现明显的问题。

多数现代系统使用的是具有相同垂直和水平维度的方形像素，但对于图像上任何方向上的相同距

离，还需要使观看方向垂直于表面。如果不垂直，那么如第 4 章中所讨论的那样，需要使图像扭曲。有些系统，尤其是那些不具有方形像素的系统，允许对水平和垂直方向建立不同的空间校准。对于基于像素计数的面积测量，这很容易处理。但对于来自边界表示的面积计算，或者对于长度测量、距离测量以及第 11 章中的形状参数测量，这种差异可以会造成严重问题。

10.9 圆和椭圆

求出面积后，将其表示为等效圆直径很方便。等效圆直径是一个线性大小测度，可由面积计算为

$$\text{等效圆直径} = \sqrt{\frac{4}{\pi} \cdot \text{面积}} \tag{10.12}$$

图 10.49 显示了几个不具有不同大小和形状的特征，并基于净面积（像素计数）给出了等效圆直径。不同形状或方向的特征会欺骗我们的眼睛，并使得我们很难判定其相对大小。等效直径值是可方便地表征特征尺寸的一个比较参数。

我们通常使用圆来表示大小，因为直径这种线性测度不会考虑形状的细节。除了使用具有相同面积的等效圆作为特征外，也可使用内切圆或外接圆来表示特征（见图 10.50）。外接圆由外接多边形的两个或三个角点确定，由这两个或三个角点确定的外接圆应包含所有其他的角点［找到这样一个圆的简明算法见（Arvo，1991）］。内切圆由特征欧氏距离图的最大值来定义，即最大像素标记圆心，其值给出半径。

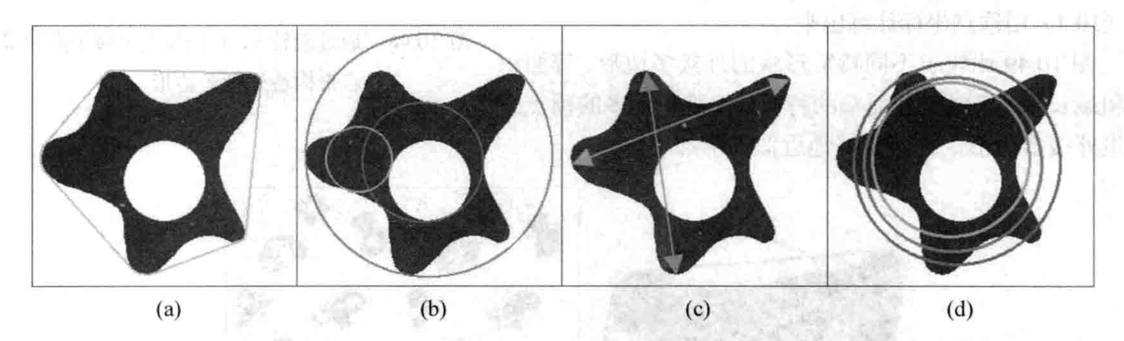


图 10.50 不同测量尺寸的一个不规则特征：原始二值特征及其凸壳（橙色）；(b)最小外接圆（红色）、最大内切圆（绿色）和忽略内部孔洞的最大内切圆（蓝色）；(c)最小（洋红色）和最大（绿色）卡尺维度（注意它们不正交）；(d)净面积（绿色）、填充的面积（红色）和凸状面积（蓝色）的等效面积圆，每个圆的圆心是相应面积的质心

由于多数真实的特征具有不规则的形状，因此很难找到能真实地描述它们并对它们进行分类和比较的大小测度。一种尝试是使用等效圆直径。识别各向并不等大的特征或近似为圆形的特征时，可使用椭圆来表示该特征。椭圆的两个轴可描述特征的大小、圆度甚至方向。

确定椭圆长轴和短轴的方法有多种。这些方法实际上表示了特征尺寸和形状的不同方面，除非用户了解并仔细对等它们，否则椭圆所提供的帮助与误导相比要少得多。

使用特征的最小和最大卡尺维度（如使用外接矩形顶点间的距离），可定义椭圆的两个轴。最大卡尺维度可很好地表示最大维度，并且至少能在搜索的步长范围（如 10°）内指出特征的方向。若将最大卡尺维度取为椭圆的长轴，则可将最小卡尺维度取为椭圆的短轴。这种方法存在几个问题。首先，对于长且窄的特征，该值可能会严重高估实际的最小尺寸。其次，最小维度方向通常并不垂直于最大维度方向（如一个正方形的对角线和边）。第三，所得椭圆的面积与特征的面积不同。

由于将椭圆的宽度基于最小卡尺维度是令人怀疑的，因此这种方法的一种改进使用了最大卡尺维度作为椭圆的长轴，并根据椭圆面积与特征面积相等来求出短轴。由于椭圆的面积是$(\pi/4)ab$，其中 a 和 b 是轴，因此一旦求出长轴，就可根据特征的面积求出短轴。由最大卡尺维度搜索中的步骤，可求出方向角的近似值，由式（10.7）中所示的矩也可计算出方向角。这样得到的椭圆看起来会更长和更窄。

由式（10.7）中的矩也可生成一个拟合椭圆。这也许是最鲁棒的测量，但它需要的计算量很大，因而很少使用。相反，许多系统不将椭圆拟合到特征区域内的所有像素，而是拟合到沿特征边界分布的像素。这一过程计算上更简单，但很难解析地证明，且沿特征边界分布的任何部分的不规则会明显使椭圆形的大小、形状和方向产生偏差。

10.10 卡尺维度

上面描述和说明的卡尺维度是特征大小的另一种描述。最大卡尺或最大 Feret 直径有时也称为特征长度，因为它是圆周上任意两点之间的最长距离。对像素或边界点排序，找到最小和最大坐标值，然后取其差值，可求出水平或垂直方向上的投射尺寸或阴影尺寸。这些尺寸是在上面介绍特征的外接矩形时引入的，可用于校正随机放置的图像视野与边缘相交的概率。

使用在不同方向确定的极值点，可构建上面所示的外接多边形。使用这样一个多边形时，通过对每对角点排列，可找到最小和最大卡尺直径。有着最大分离距离的一对顶点可近似逼近特征的真实最大尺寸。当实际最大弦正好是角度步长之间的一半时，会出现最大误差，此时由

$$测量长度 = 真实长度 \cdot \cos(\alpha/2) \tag{10.13}$$

测量出的长度会较短。对于上面提到的 10° 步长的例子，5° 的余弦是 0.996。这意味着测量误差比 0.5% 还低。对于实际最大尺寸是 250 像素的特征，旋转坐标轴所得到的该值，在最坏情形的方向上要短一个像素。

如果使用相同的方法来求特征的最小卡尺维度（有时也称为宽度），则它等同于找到外接多边形上靠得接近的一对相反的顶点。此时误差要远大于最大卡尺维度的误差，因为它取决于角度的正弦值和特征的长度，而非特征的实际宽度。长度为 L、实际宽度为 W、方向距最近旋转步长同样是 5° 的一个较窄特征，其估计宽度为 $L\sin(5°) = 0.087L$。这甚至并不取决于宽度 W，且如果实际长度 L 大而宽度 W 小，那么会存在明显的误差。

图 10.51 显示了这样一种情形，其中测量的长度是合适的，它可按上面的说明求出。样品是平板扫描仪上散布的米粒。美国农业部（USDA）关于长米粒的标准规定其最长部分不短于 6 毫米。测出这些米粒的长度分布，可快速得出结论。

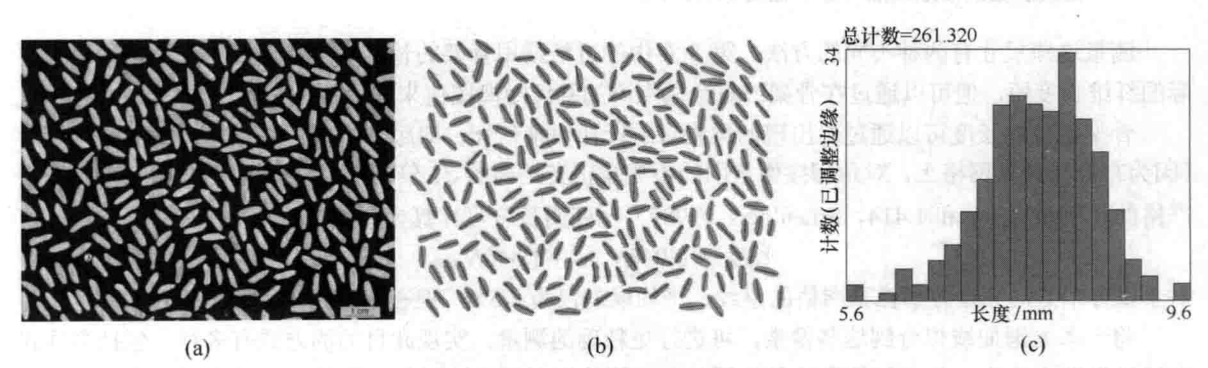


图 10.51　特征长度测量：(a)原始图像（米粒）；(b)经阈值处理后的二值图像，其上叠加了米粒长度；(c)米粒长度分布图

当然，对边界点排序，找到具有最大分离距离的两个边界点，也可直接搜索最大弦。加速这种搜索的方法很多，否则就必须计算所有边界点对之间 x 坐标之差和 y 坐标之差的平方和。一种方法是只考虑位于等效圆外的那些点。求出面积（此时是包含有任何内部孔洞的填充后的面积）并用式（10.12）将其转换为等效直径，可得到圆的尺寸。将圆心放在特征质心上，可覆盖圆内的所有点而只留下离得较远的那些点。这些点就是离得最远的候选点对。

由外接矩形大小，可更严格地限制对待搜索点的选取。如图 10.52 所示，与外接矩形的每条边都正切的弧线所覆盖的区域，包含了特征的大部分边界。搜索中只须考虑弧线上或弧线外的那些点，且一条弧线之外的那些点只需与对面弧线之外的那些点结合。对于边界上可能包含上千个点的复杂特征，在需要精确的最大尺寸时，这些选取算法的优点与所需计算量相比，就非常值得。

求直物体或弯物体宽度的一种方法是，使用骨架对欧氏距离图进行抽样来得到沿中线分布的值。如第 8 章所示，它们是沿特征长度分布的每个点处的内切圆的半径。取决于标本的性质和测量目的，可使用这些值的最小值、平均值或最大值作为测量的宽度。宽度值的标准差也可表征宽度的变化。

考虑图 10.53 所示的 S 形状特征。如果它是一个刚体，如铁链上的铸铁钩，则应按最小和最大卡尺维度来定义长度和宽度。另一方面，如果物体是一条蠕虫或一根弹性纤维，则应测量纤维的轴向长度和横向宽度。为了区分与通常称为长度和宽度的最大和最小卡尺维度，它们有时称为纤维长度和纤维宽度。在特特殊情形下，最终由用户确定哪些参数可以有效地描述特征。

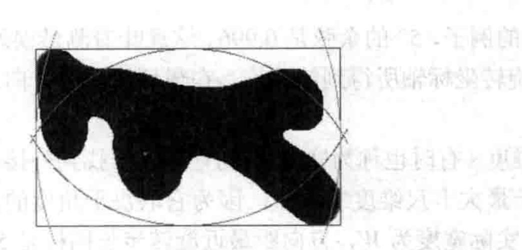


图 10.52　限定边界点搜索范围所构建的最大弦。分别以外接矩形水平边缘和垂直边缘的中点为圆心，画出了的蓝色弧线和绿色弧线。只有特征边缘上的红点不在弧线范围内，这些点是最长弦的候选端点（显示在图 10.49 中）

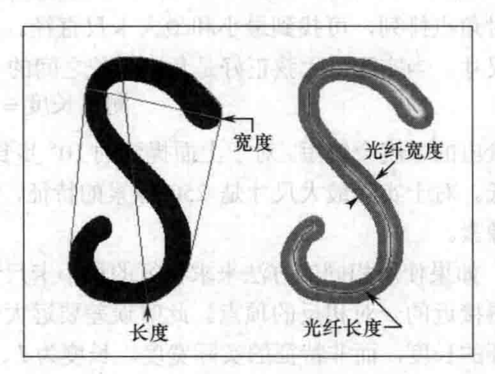


图 10.53　卡尺维度（长度和宽度）与由骨架和欧氏距离图（光纤长度和光纤宽度）求出的值的比较

测量这些尺寸有两种不同的方法。第 8 章中的方法使用骨架的长度来测量纤维的长度。骨架比实际的纤维长度短，但可以通过在骨架的每个端点添加欧氏距离图值来进行校正。

骨架本身的长度可以通过对出现的像素对进行计数来估计，即跟踪那些边缘和角点相接的像素对（因为在方形像素网格上，对角相接像素间的距离是正交距离的 $\sqrt{2}$ 倍）。这会稍微过估长度。替代采用严格的几何距离 1.0 和 1.414，Smeulders（1989）已证明长度可计算为

$$长度 = 0.948N_{正交} + 1.340N_{对角} \tag{10.14}$$

对于以各个方向穿过方形像素网格的直线，平均误差仅为 2.5%，但在特定方向上会出现较大的误差。

将一条光滑曲线拟合到这些像素，可进行更精确的测量。实现此目的的方式有多种，包括多项式曲线和贝塞尔曲线。最有效的方法是使用下面介绍的用于周长测量的相同技术，其中对图像进行了平滑，且由这些像素内插了一条超分辨率直线。

对纤维形状做出几何假设，是估计纤维长度和宽度的一种不太准确（但仍在使用）的老方法。如图 10.54 所示，若假设特征的长度为 F（纤维长度），均匀宽度为 W（纤维宽度），则特征的面积（A）和周长（P）分别是 $A = FW$ 和 $P = 2(F + W)$。面积和周长可以直接测量。如下所述，周长更难以求出，但对于此处假设的具有平滑边界的条带，可以准确地测量出周长。然后，由测出的周长和面积，按如下公式求出纤维的长度和宽度：

$$F = \frac{P - \sqrt{P^2 - 16A}}{4}, \quad W = \frac{A}{F} \tag{10.15}$$

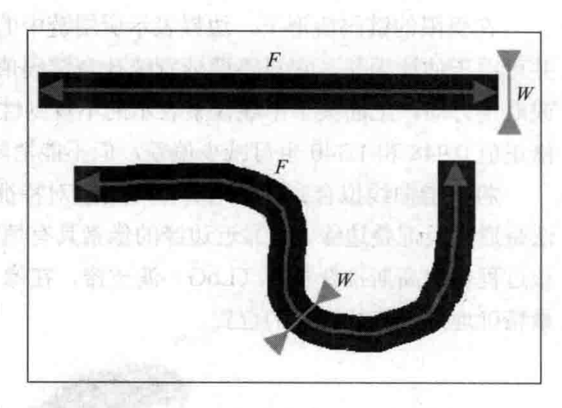


图 10.54 具有相同宽度和方形末端纤维的几何模型。根据顶部显示的矩形，可由测量的面积和周长计算出长度 F 和宽度 W，对于弯曲形状也基本正确

对该模型可稍做修改，如假设条带的两端是圆形而不是方形，但原理不变。这种方法的困难之处在于，它对所用形状模型和实际纤维形状之间的差异的敏感性，以及与周长测量相关的一些问题。例如，如果特征的宽度不一致，或特征具有分支，那么由骨架和 EDM 得到的平均宽度仍然是一致且有意义的结果，而由式（10.15）计算得到的平均宽度就不是如此。

10.11 周长

特征的周长是我们熟悉的一个几何参数，它在连续实空间中能很好地定义，但在像素化图像中就要差一些。测量出能正确描述物体的值并不简单。最终，测量取决于是将像素视为一个点还是视为一个方形（Hetzner，2008）。有些系统通过统计那些与背景接触的像素来估计围绕物体的边界长度。这会低估实际的周长，因为如上面对骨架长度的测量所述，角接触像素之间的距离大于边接触像素之间的距离。这种误差取决于特征方向，而且像方形这样一个简单物体的周长会随相机的旋转而变化。图 10.55 说明了当一个近似恒定大小的正方形旋转时，通过统计像素边缘得到的周长，与由链码得到的更准确的数值进行比较所出现的周长变化。测量值对图像中相同特征变化方向的敏感性，经常用于测试系统性能。

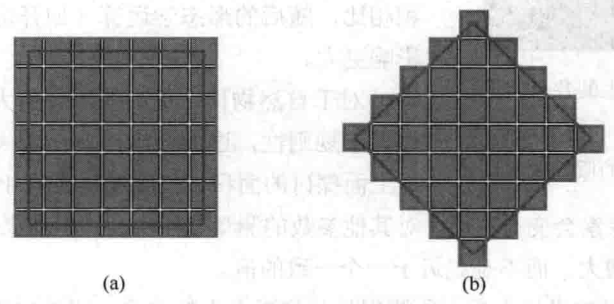


图 10.55 链码所得周长与像素边缘长度之和所得周长的比较，它与方向有关：(a)外部边缘之和（红色）为 28 个单位，中心到中心链码距离之和（蓝色）为 24 个单位；(b)外部边缘距离（红色）为 30 个单位，中心到中心链码距离之和（蓝色）为 $20\sqrt{2}$ 即 28.28 个单位

使用边界表示来表示特征时，对连续点之间的毕达哥拉斯距离求和可估计出周长，即

$$周长 = \sum_i \sqrt{\left(x_i + x_{i-1}\right)^2 + \left(y_i - y_{i-1}\right)^2} \tag{10.16}$$

在受限的链码情形下，边界表示所用链中的链接长度要么是 1.0 个像素，要么是 1.4142 个像素，并可用于估计周长。它只需要分别统计奇链码值的数量和偶链码值的数量，因为它们能区分正交方向或对角方向。上面关于中线像素表示的不规则性探讨，同样适用于边界线，且使用来自式（10.14）的修正值 0.948 和 1.340 也可减少偏差，但不能消除偏差。

将平滑曲线拟合到特征边界，可得到对特征边缘方向最不敏感的最准确的周长。这样做的一种方法是通过反混叠边缘，使靠近边缘的像素具有填充阶梯的灰度值（Neal et al.，1998），如图 10.56 所示。该过程使用高斯拉普拉斯（LoG）滤波器，在像 Canny 边缘定位器这样的算子中使用这种滤波器，可最精确地定义特征边界的位置。

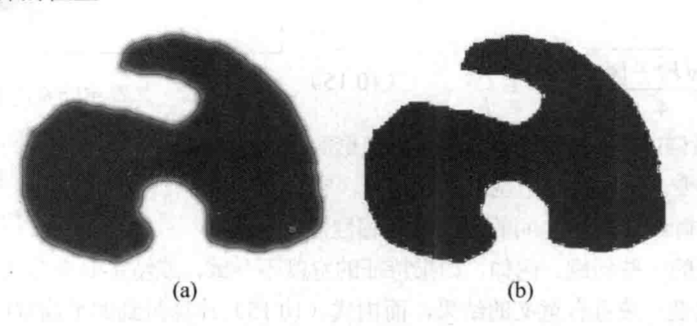


图 10.56　为周长测量拟合一条平滑的边界：(a)平滑（反混叠）特征上的轮廓线；(b)叠加到原始像素上的轮廓线

画出并测量跨越反混叠像素的直线段，就构建了实际的边界。这些线段是在（沿根据灰度值线性内插的像素边缘分布的）各点之间画出的。这些点对应于中点值所在的位置，它将各个像素当做间隔点处理。结果等同于使用超分辨率和双线性内值来填充新的较小像素而增大图像，如图 10.57 所示。

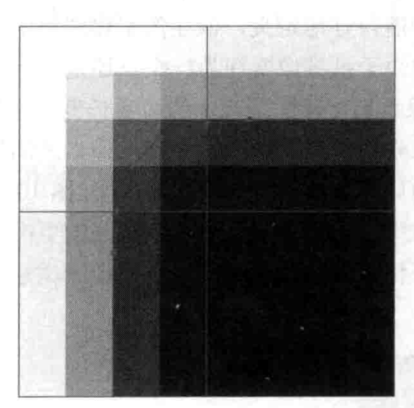


图 10.57　在平滑像素边缘上的各点间进行边界内插，显示了锐角已变圆。蓝线表示内插前的原始像素尺寸

尽管超分辨率边界测量技术对特征旋转是准确的和鲁棒的，但周长测量的基本困难是，对于多数物体，周长本身的定义不清，且依赖于放在倍数。沿特征边界分布的像素通常具有界于物体和背景之间的亮度或色彩值，因为检测器的有限区域重叠了不同角度的边界。图像处理和阈值处理如何更好地或一致地识别边界位置，限制了任何后续的测量。与面积相比，随后的形态学运算（如开运算和闭运算）对周长的影响更大。

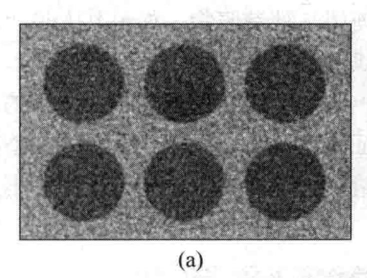
对于自然物体，更高的图像放大率通常会揭示出更多的边界不规则性，进而会给出一个较大的周长值。但这不适用于上面探讨的面积、长度或其他尺寸。随着成像尺度的改变，与特征大小相比，各个像素会变得更小，对其他参数的测量当然也会，但结果会收敛到一个最佳估计。对于周长，该值通常会增大，而不是趋近于一个一致的值。

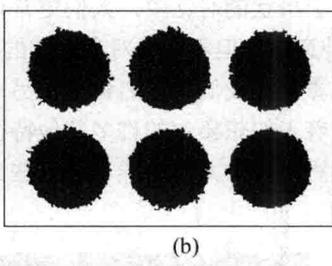
在许多情形下，在对数坐标上画出所测周长与像素大小的关系，或画出所测周长与测量分辨率的关系，可得到一条直线。这是一种表示分形行为的自相似性，则直线的斜率可得出边界的分形维度。第 11 章中将使用它来测量特征的形状。

即使特征边界是不严格的分形（即由周长与测量尺度关系的线性对数图表示的自相似性），增大成像放大率，通常也会增大测量的周长。唯一的例外是平滑（欧几里得）的物体，如细胞膜或表面，其中的张力或能量项因为物理原因产生了局部平滑。这种对分辨率的依赖使得周长值疑似为物体的真实

描绘子，并至少部分地受到成像方法和尺度的影响。

此外，图像中的任何噪声都会导致边界变粗并增大周长。图 10.58 显示了一个例子。随机灰度噪声叠加在 6 个相同的圆上。这些圆的直径为 80 像素，测量面积和周长分别为 5024 像素和 254 像素（分别接近计算出的理论值 5026.5 和 251.3）。测量一系列带噪图像后，随着像素加入特征或从特征中删除，对得到的面积求平均来校正均值。但周长总是增加的，如表 10.1 中的结果所示。面积测量值的变化只有 0.5%，周长测量的变化超过 5%，且平均值太高。成像过程中由噪声引入的这种偏差是周长测量中要考虑的另一个因素。使用形态学开运算或闭运算平滑边界，可以减少周长值，如图所示，但这假定了事先已知边界应被平滑。


(a)


(b)

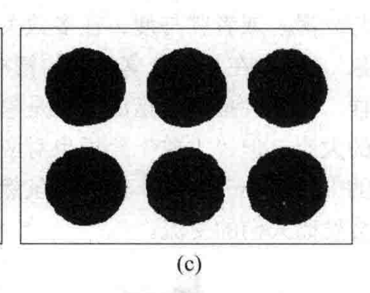

(c)

图 10.58　噪声对测量的影响：(a)带有加性灰度噪声的 6 个相同的圆；(b)经阈值处理并填充内部孔洞后的圆；(c)图(b)经平滑和基于 EDM 的开运算（半径为 6 像素）后的圆。表 10.1 列出了面积和周长的测量数据

表 10.1　图 10.58 中圆的测量结果

圆	经阈值处理后的面积	经阈值处理后的周长	应用开运算后的面积	应用开运算后的周长
1	5027	365.36	4997	275.93
2	5042	344.53	5041	249.57
3	4988	350.08	4996	274.44
4	5065	355.92	5040	252.35
5	5020	375.76	4994	279.46
6	5030	321.60	5051	249.56
平均	5028.67	352.21	5019.83	263.55
标准差	25.42	18.66	26.77	14.43
原始值	5024	254		

测量周长时，同样存在面积没量时的三种选择。总周长包括特征内的围绕任何孔洞的边界长度。外部周长、净周长或填充的周长则不包括内部孔洞。凸周长是凸壳或外接多边形的长度，以及连接特征边缘凹口的桥。它对噪声或图像分辨率不太敏感，因为它不跟踪特征边界的所有点。凸周长可用式（10.16）计算为外接多边形顶点间的毕达哥拉斯距离之和。在特殊应用中，针对面积测量的考虑，同样适用于周长。

第11章 形状表征

11.1 描述形状

人类语言并不能很好地描述形状。描述形状的形容词很少，即使有，也只是一些定性的描述，如粗糙与光滑，或者胖与瘦。在多数关于形状的对话中，人们更常使用某种参照物。如果某人说“形状像大象”，则是在与你分享相同的精神意象。但不同人对参照物的理解是不同的。如图 11.1 所示，“大象一样”的定性描述所指的似乎是象鼻。当我们将该名称用于另一物种如鼩鼱或海豹时，所指的应是鼻子的大小。但“大象”一词也对应着人们想象中的许多其他特征，如大小、象牙或耳朵（如迪斯尼公司的卡通“小飞象”）。因此，虽然短语“像大象一样”在某些情形下可方便于人们的沟通，但有时候却会妨碍人们的交流。

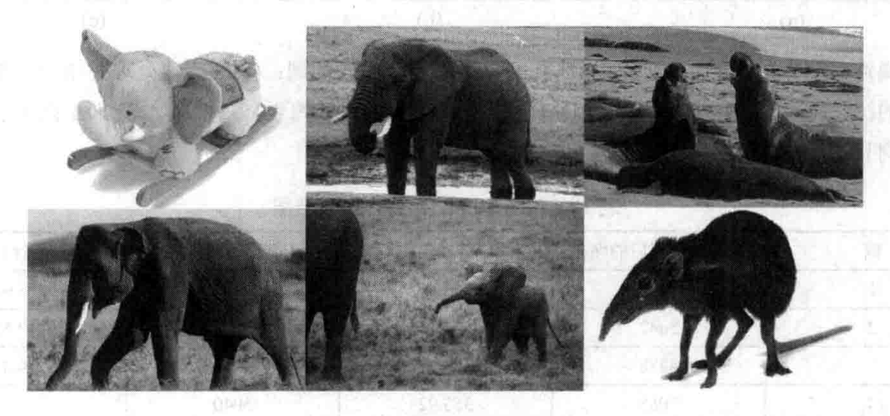


图 11.1 大象（从左上方起）：小孩的摇摆玩具，非洲象，海象，亚洲象，小象，象鼩

美国女画家 Georgia O’Keefe 曾说过：“我只能用颜色和形状而非其他方式来表达事物。”

有些词似乎是描述形状的形容词，但仔细思考后却发现不是。“圆的”一词，我们通常理解为“像一个圆”，或者在三维情形下理解为“像一个球”。它还可理解为一个圆筒或一个圆环，具体取决于情形和视角，有时可用于区分直线边界和弯曲边界及凸状边界。图 11.2 给出了几种形状，这些形状在不同情形下可以用“圆的”来描述。当将形状理解为“像一个圆”时，“圆”是一个名词。下面将要介绍的许多简单测量，所测的是二维形状偏离圆的程度，并用名称“圆度”表示。遗憾的是，不像圆的方式有多种，如将圆拉长形成椭圆、将边压扁形成多边形等。

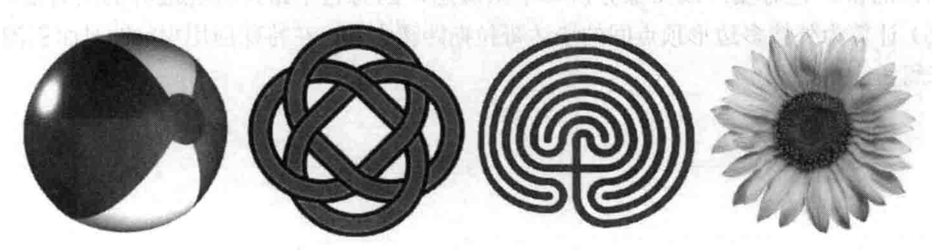


图 11.2 单词“圆的”表达的 4 种形状（左起）：沙滩排球，凯尔特结，克里特岛迷宫，向日葵

圆的几何定义很简单：与中心等距的点的轨迹。另一个定义是：各个方向上直径相同的一种形状。

在许多方向上测量某个特征的半径，可简单识别该特征的形状是否为圆形。直觉有时并不代表事实。例如，英国发行了 50 便士和 20 便士面额的几种硬币，它们的形状是等直径曲边七边形（见图 11.3）。曲边造成所有方向上的直径相同，因此可在自动售货机上使用。奇数个顶点的任何多边形都可采用这种方式。



图 11.3　七边形等直径硬币

确定工业制品（如汽车发动机汽缸）是否真正为圆形的高精度测量仪器必须避免落入这种陷阱，即相对简单的测量结果看起来与期望的结果相符，但实际上却不能提供所需的信息。测量二维或三维形状来确定它是否为“圆的”并不简单。实际形状只是“近似”为圆形时，简明而有意义地描述它并不容易。

形状在人类视觉中最重要的用途之一是识别。形状通常比大小更重要。识别靠近你的老虎比识别它是一只大老虎还是一只小老虎更重要。在采集用于晚餐的植物时，形状可以区分可食用植物和杂草。

计算机视觉的重要目的之一是找到形状的数值描述。可以分析或匹配数字的软件是有效且简单的。选取提供足够独特性且同时能让人们表征形状的一组最佳数字并不容易，因为数字和我们的日常经验并无多大的对应性，并且参数都具有抽象的“人工”特性。当基于人们的主观检查来指定形状分类的形容性描述时，找到测量过程来产生与这些标记相符的结果并自动地让计算机进行测量是很困难的。例如，Les & Les（2005）提出了像“薄的，凸状，圆形，复杂的”这样的分类，但并未指出定义这些组合的方式。

11.2　无量纲比

最早的一类数值形状描述子是尺寸参数的组合，这种组合抵消了量纲。例如，“长度/宽度”给出了长宽比，改变特征的尺寸不会改变长宽比的数值。当然，如第 10 章所述，要假设正确测量了长度和宽度。

由于尺寸有几十种，因而可用这些尺寸的数百种组合来作为无量纲的形状描述子。常见的组合只有几种，但即使这些组合也会受命名约定不一致的困扰。表 11.1 总结了本书中尺寸测量所用的一些常

表 11.1　典型的形状描述子

形状因子 $=\dfrac{4\pi\cdot\text{面积}}{\text{周长}^2}$	圆度 $=\dfrac{4\cdot\text{面积}}{\pi\cdot\text{最大直径}^2}$	长宽比 $=\dfrac{\text{最大卡尺维度}}{\text{最小卡尺维度}}$
延伸率 $=\dfrac{\text{纤维长度}}{\text{纤维宽度}}$	曲度 $=\dfrac{\text{长度}}{\text{纤维长度}}$	凸度 $=\dfrac{\text{凸面周长}}{\text{周长}}$
实心度 $=\dfrac{\text{面积}}{\text{凸面面积}}$	致密度 $=\dfrac{\sqrt{\left(\frac{4}{\pi}\right)\text{面积}}}{\text{最大直径}}$	修正率 $=\dfrac{\text{内切圆直径}}{\text{外接圆直径}}$
范围 $=\dfrac{\text{净面积}}{\text{外接矩形}}$		

用形状参数。有些系统会将这些参数定义为所示公式的倒数，或者会省略常数乘数（如 π）。此外，在一些特殊系统中，相同的计算可能会被一些相当不同的名称调用，或者为不同的关系使用表中所示的名称。40 年前的早期图像分析系统中，就已引入了本文中称为形状因子的常用形状描述子，它在 Photoshop 中称为“圆度值”，而在 Image Pro Plus 中则称为“圆度”。

有些描述子在某些特殊领域的应用非常成功。例如，修正率（内切圆与外接圆的比率，也称为半径比）用于度量人造纤维的截面，如编织地毯所用的纤维。若这些“星状”截面的内切直径增大，则表明纺嘴有磨损，且纤维的弹性降低、质量增大。计算机测量出现之前，这个比值是使用圆形模板来视觉拟合截面求出的。

当然，由于测量的目的是使得比较更为容易或使得比较更为量化，因此用户的责任就是要确保任何特定形状的描述子的含义容易理解，因为所选择的这种描述子反映了特征形状的变化关系。由于形状因子的应用如此广泛，因此一定要了解它测量的是什么。图 11.4 中显示了面积不变而周长增大的情况。所有这些特征几乎都是各向等大的，长宽比都接近于 1。图 11.5 中的形状显示了拉伸以增大长宽比而同时保持面积不变的过程。这些形状的形状因子变化不大。

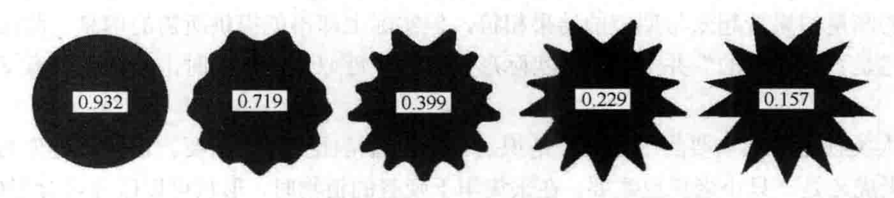


图 11.4　形状因子：这些形状具有相同的面积，如图所示，增大周长会改变形状因子的测量值

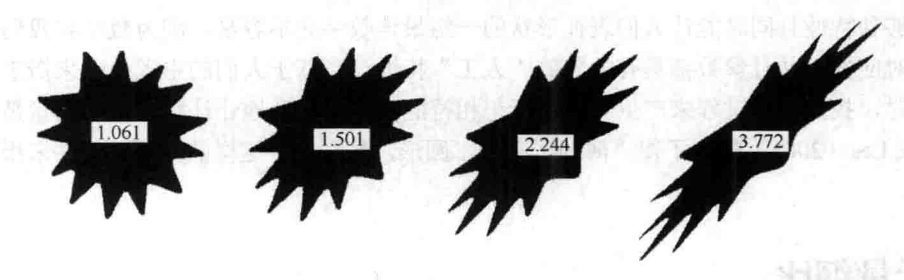


图 11.5　长宽比：这些形状具有相同的面积，如图所示，形状因子的较小变化会导致长宽比的很大变化

这两幅图形表明的事实是，虽然只有一种方式“像一个圆”，但却有很多偏离圆的方式。随着特征以各种方式拉长或生成不规则边界，各个无量纲形状参数会表现得不同。对于一个理想的圆，形状因子、圆度和长宽比都恰好为 1.0。在真实的像素网格上，特征只能近似为圆，具体要取决于测量周长所用的方法，而形状因子小于 1。图 11.6 通过使用已在第 10 章中描述的超分辨率边界测量方法显示了这种情况；尽管有形状内在的像素化逼近，但对于直径大于 10 个像素的特征，形状因子值相当一致（但小于 1.0 这个理想值）。

使用形状因子（或任何无量纲的比值）来表征形状的最大限制既不是测量精度，也不是它与人们判定形状的方法不同。相反，是因为这些参数都不是唯一的。图 11.7 显示了 12 个形状看起来完全不同的特征，但它们具有相同的形状因子值。注意，无量纲比只获取关于形状的一种或非常有限的信息。

其他形状参数代表了形状的不同特性。例如，图 11.8 显示了具有不同曲度值的几个形状。图 11.9 以一种不同的方式说明了这些参数之间的差异。“特征”是以一种衬线字体（Times Roman）印刷的 26 个大写字母。在每一行，对不同的形状参数测量进行了颜色编码。每组中的颜色变化是从红色（最大数值）到品红色（最小数值）。每种形状因子的独立变化非常明显。这一方面表明了形状因子是识别特征的强大工具，但另一方面，由于这样的因子繁多且人类视觉无法对它们进行分类或估计，因此并不使用这些参数来进行特征识别。如果使用另一种字体，特别是无衬线字体，那么对各种形状测量的相对值是不同的。

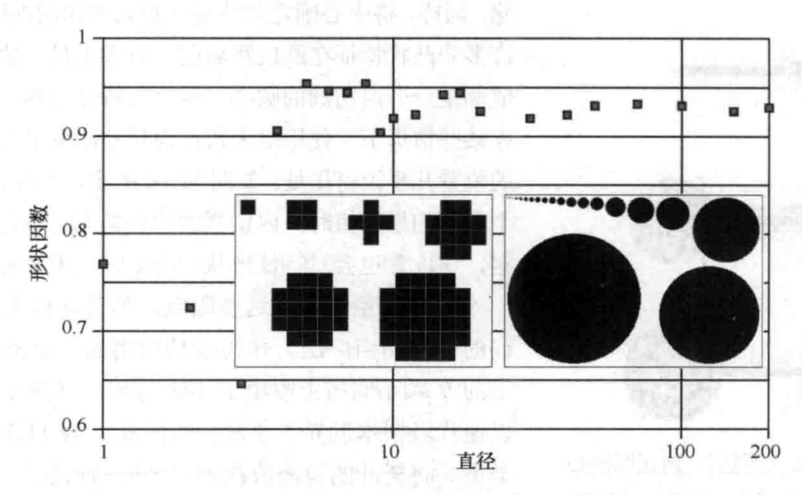


图 11.6　随圆直径变化的形状因子（已放大最小的圆来显示形成该圆的精确像素）

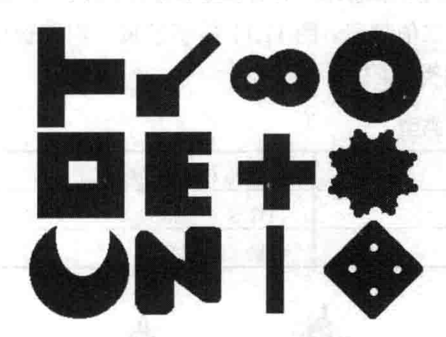


图 11.7　一组视觉上可以区分但有相同形状因子值（0.44）的特征

0.928　0.582
0.300　0.411

图 11.8　具有不同曲度值的 4 个特征

ABCDEFGHIJKLMNOPQRSTUVWXYZ 形状因数
ABCDEFGHIJKLMNOPQRSTUVWXYZ 圆
ABCDEFGHIJKLMNOPQRSTUVWXYZ 长宽比
ABCDEFGHIJKLMNOPQRSTUVWXYZ 伸长率
ABCDEFGHIJKLMNOPQRSTUVWXYZ 凸度
ABCDEFGHIJKLMNOPQRSTUVWXYZ 坚固性
ABCDEFGHIJKLMNOPQRSTUVWXYZ 紧密度
ABCDEFGHIJKLMNOPQRSTUVWXYZ 广度
ABCDEFGHIJKLMNOPQRSTUVWXYZ 卷曲度

图 11.9　一系列形状（字母表中的字母）的测量。在每一行中，对应于所示的描述子，特征的颜色编码为不同形状参数的相对数值（红色表示高，品红表示低）

根据第 10 章中介绍的各种位置定义，可计算出另一组简单的无量纲形状参数。对于一个完美对称的特征，如圆形或规则多边形，其质心和几何中心（外接圆的圆心）重合。如果特征不对称，就要将两点间的距离除以外接圆的半径来归一化，得到 0 和 1 之间的一个值来测量不对称的情况。未加权的质心（其中所有特征像素都同等地计数）和加权质心（其中的像素根据某个校准密度函数来计数）间的距离可以按同样的方式使用。

使用其他位置测量的类似比值用处不大。例如，将中心确定为周长像素的平均位置对分辨率非常敏

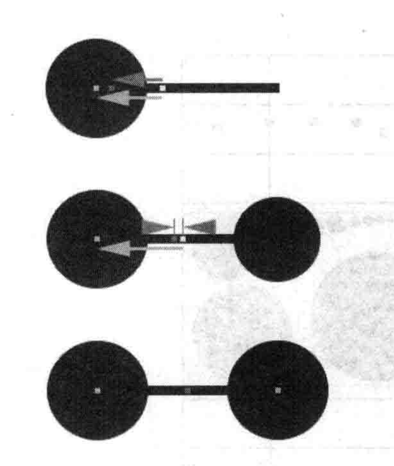


图 11.10 几何中心（黄色）、内切圆圆心（青色）和质心（红色）的位置，显示了不对称性和形状的效果

感。同样，将中心确定为特征中最大内切圆的圆心也会有问题。许多非凸状特征在欧氏距离图上有多个最大值。每个局部最大值都是一个内切圆的圆心（至少与特征边界上的 3 个点相切）。在某些情形下，使用最大值作为尺寸测量是有意义的，但中心的位置几乎没有用处。如图 11.10 所示，当特征在 EDM 中有多个相等的最大值时，内切圆的圆心是无法确定的，即使可以确定，该位置也会对特征形状的许多方面不敏感。

尽管无量纲比有这些限制，但其计算上的简单性使得它们的应用相当广泛。作为成功应用的一个例子，我曾从院子中的 9 颗不同树上收集了 148 片树叶（通常来自同一分枝，以便代表树木世界中非常少的样本）。表 11.2 中列出了树木的类型。将树叶的背面放在台式平板扫描仪上，以 200 像素/英寸的分辨率进行了扫描，扫描后的图像经阈值处理后生成了可进行测量的二值表示。图 11.11 显示了每一组图像中的一个示例树叶。注意，人类视觉可很容易地识别出这些不同类型的树。

表 11.2 树木类型

樱桃（樱桃属“关山樱”）	山茱萸（北美花茱萸）	山核桃（山核桃属油丹）
桑椹（桑属）	红橡木（南美红栎）	白橡木（白橡树）
香枫（北美枫香）	银枫（槭属）	红枫（红花槭）

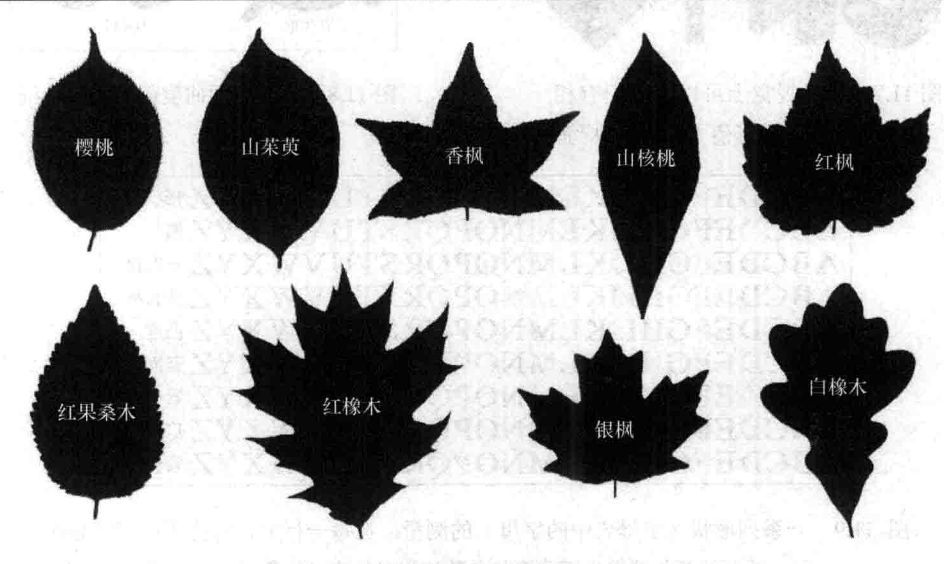


图 11.11 树叶形状示例（比例尺不同）

对每片树叶计算上面所示的无量纲比参数，并用 SAS JMP- 8 软件包分析数据集来进行判别分析。分析是逐步进行的，以便在没有误分类的前提下，确定可对树叶正确分组的最少重要项。图 11.12 表明，只用三个参数（圆度、实心性和凸度）即可正确地鉴定每片树叶。

与自动识别植物叶片的更大雄心相比，这只不过是一个小小的说明性例子。Belhumeur et al.（2008）描述了 LeafView 系统，该系统基于平板电脑（PDA 或智能手机），可以获取叶片图像、处理图像并进行必要的测量，访问数百种植物的数据库，并向用户给出最匹配的 10 幅图像，但与植物的在线数据库相比，这个系统的功能仍然相当有限。这样的在线数据库包括美国国家植物标本馆的 9 万种标本、纽

约植物园的 12 万幅图像、英国皇家植物园丘园的 35000 幅图像、密苏里植物园的 35000 幅图像，以及世界各地的其他资源。能与这种大型数据集进行交流的专家系统将彻底改变传统的“野外生存指南”。

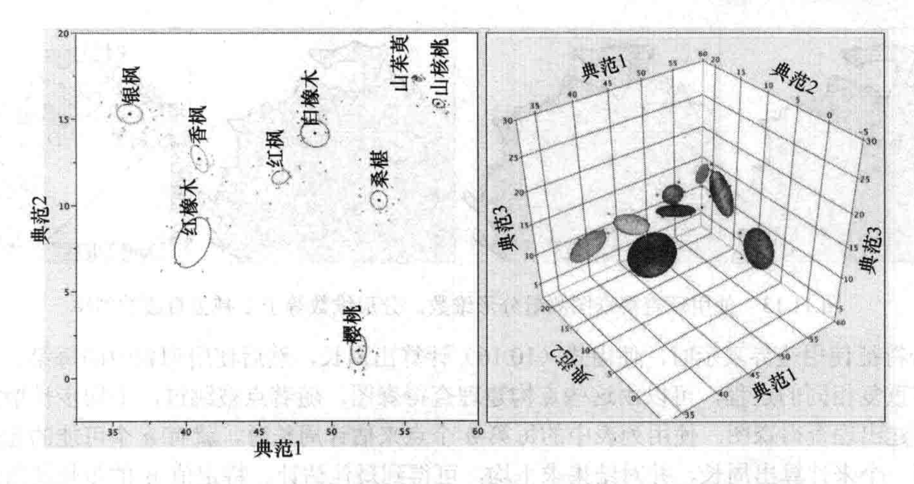


图 11.12　使用三个经典的尺寸和形状测度即可判断树叶样本的分析结果

第 12 章中将详细探讨物体的分类和识别过程。这里的目的是要表明，无量纲比参数在许多实际情形下会是有效的描述子。

11.3　分形维数

上面给出的几个形状参数中包含特征的周长。如第 10 章中指出的那样，这通常是一种有问题的测度，即所得测量值会受图像分辨率和放大倍数的影响。事实上，对于许多真实的物体而言，周长这一概念可能存在根本性的缺陷。若真实世界的物体形状实际上是分形的，则周长就无法确定。在较高的放大倍数下测量物体的边界时，所得的周长总是较长。例如，考虑一片云彩：围绕云彩投影图像的边界长度是多少？测量涵盖了多个层级，从由可见光观察到的一小片云彩，到由雷达或气象卫星观察到的整个风暴系统，都表明云彩的周长服从分形几何。假定这种趋势继续到越来越小的尺度，直到水分子大小。那么，周长真正表示的是什么呢？它是物体的真实特性，还是测量尺度的人为假象？

分形物体看起来并不是自然界中的例外情形。欧几里得几何，因其良好的定义和数学上易于处理的平面和曲面，通常只被人类在较窄的维数范围内用做一种近似，或在某个能量或力占主导地位的情形（如气泡的表面张力）下用做一种近似。仅在人类感知和控制的维数或尺度内，我们才能说道路、建筑物和纸张的表面是平面、直的和欧几里得的。放大纸张表面，会发现它变得粗糙。从太空看道路，它们不再是直线。使用分形维数来描述欧氏空间中的线和面，是概念上相对较新的一种突破，它是描述粗糙度和形状的一个新工具。

总之，分形维数是物体周长（或三维情形下的表面积）随着测量尺度的减小而增大的速率。测量分形维数的方式有多种。有些方式是物理的，例如，黏附在单层表面上的气体分子的数量与分子的大小有关（分子越小，则表明小表面的不规则性越大，进而表明更多的表面积）。有些方式在频率空间更易于使用，如第 15 章中所示的检查二维傅里叶变换的功率谱。其他方式是对特征经阈值处理后的二值图像进行测量。

最广为人知的分形测量工具也许是理查得森图。最初引入理查得森图的目的是手动测量地图。这种测量方法的步骤是，首先用圆规量出某个已知距离，然后从边界上的某个点开始，使用圆规循周长跨步。步数乘以步幅长度就可得出周长。随着步长的减小，路线会跟随更多的局部不规则边界，随之所测量的周长会增加，如图 11.13 所示。画在对数坐标上，结果是一条直线，由直线的斜率可得出分形

维数。在每个端点处会出现线性偏离：在步长较大时，可能会完全错过边界；而在步长较小时，地图有限的分辨率会限制这一测量。

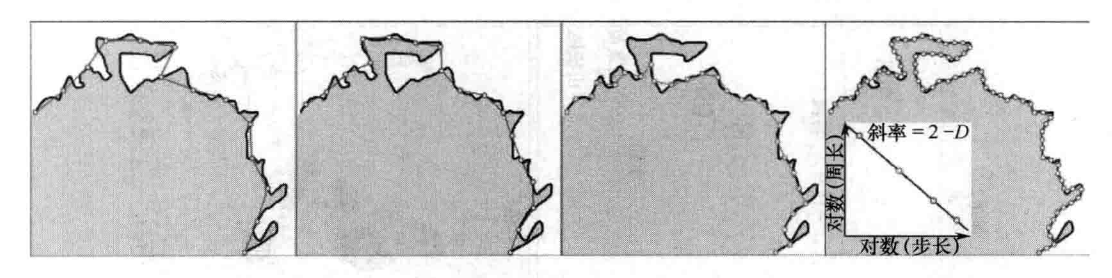
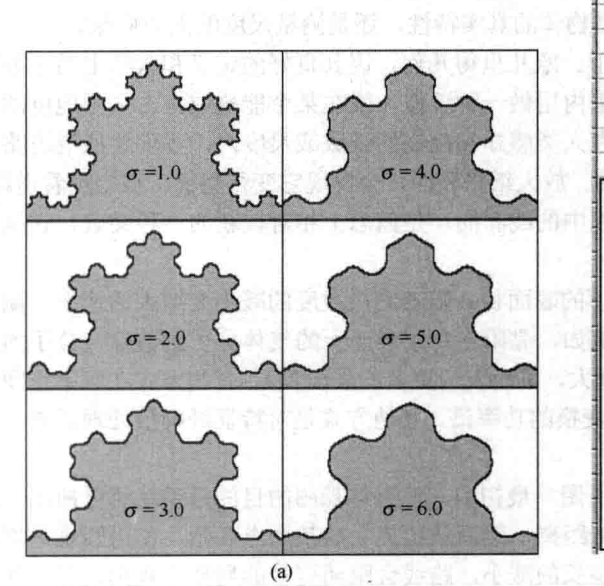


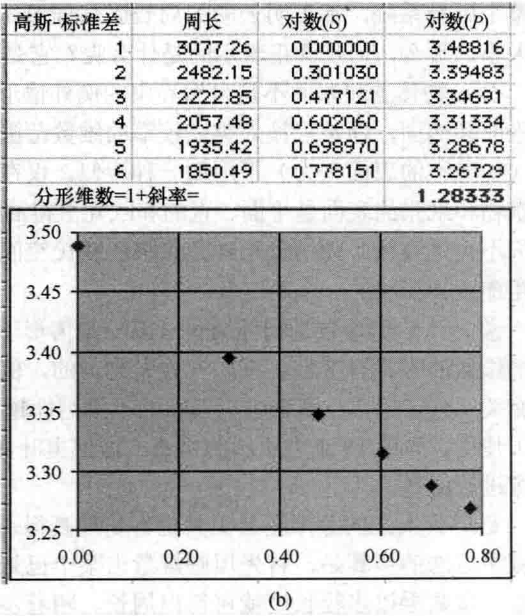
图 11.13　使用理查得森图测量分形维数。分形维数等于 2 减去直线的斜率

为某个特征使用边界表示时，使用式（10.16）计算出周长，然后使用列表中的每第二个点、每第三个点……重复相同的过程，可以由这些点构建理查得森图。随着点被跳过，平均步长增加而周长减小，进而构建出理查得森图。使用列表中的每第 n 个点来估计周长时，就有 n 个可能的起始位置。用它们中的每一个来计算出周长，并对结果求平均，可得到最佳估计。特定值 n 的步长通常由总周长除以步数得出。在所有这些过程中，我们会意识到路径的末端可能需要分数步幅，而这已包含在处理中。

当边界点的 x、y 坐标数字化成网格时，就引入了步长的变化。与圆规沿地图边界跨步的人工过程不同的是，在离最后一个点一个跨步的位置处，可能不会记录一个点，因此需要由实际的数据内插，或使用另一个点并改变步长。内插默认假设边界局部上是直线，而这与整个分形模型相悖。改变步长会使理查得森图出现偏差。结合周长测量过程所固有的问题，我们会发现经典的步长方法并不适用于数字化像素图像的测量。

但使用超分辨率周长测量方法，可以采用另一种与理查得森技术密切相关的方法。应用具有逐渐变大的高斯标准差的平滑核来模糊边界，会消除一些小的不规则边界，进而减少周长。在对数坐标系中，画出测量周长与平滑核标准差的关系曲线，得到一条直线，由该直线的斜率可得出特征的分形维数，如图 11.14 所示。维数值大于 1（一条直线的拓扑或欧几里得维数）而小于 2（平面维数），可将它视为直线向外扩展到平面的程度，或将它视为直线位于点的哪一侧的不确定性。



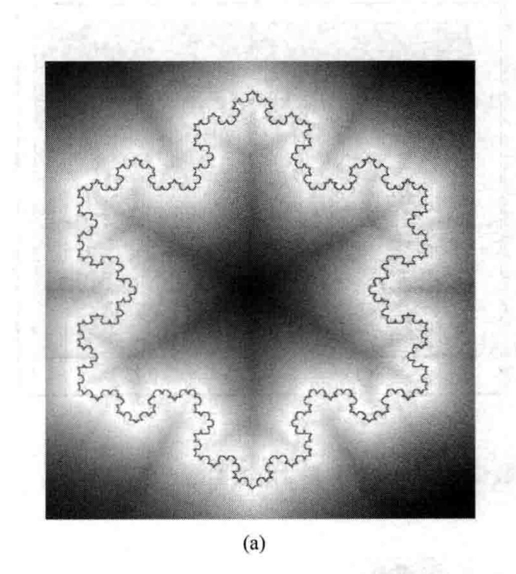


高斯-标准差	周长	对数(S)	对数(P)
1	3077.26	0.000000	3.48816
2	2482.15	0.301030	3.39483
3	2222.85	0.477121	3.34691
4	2057.48	0.602060	3.31334
5	1935.42	0.698970	3.28678
6	1850.49	0.778151	3.26729
分形维数=1+斜率=			1.28333

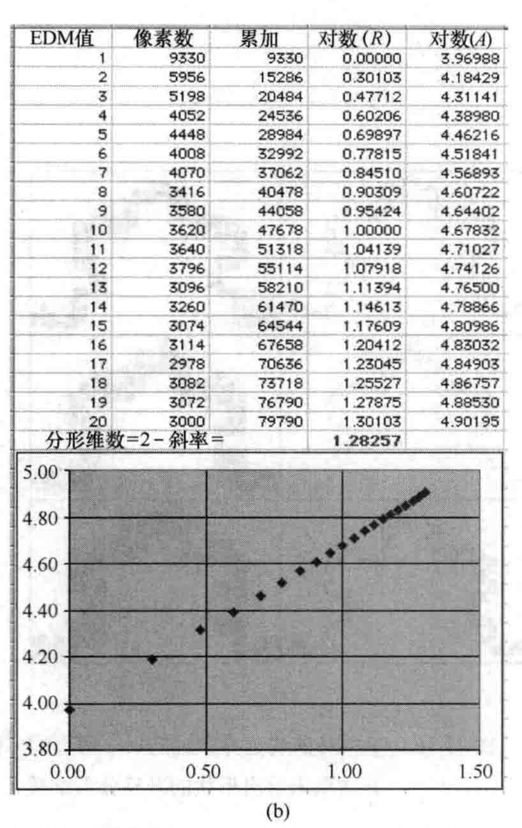
(a)　　(b)

图 11.14　(a)科赫曲线轮廓的渐进高斯平滑（理论分形维数为 1.262）；(b)周长与分形维数关系图

第二种测量技术在第 8 章中介绍二值图像处理时描述过。对边界进行不同程度的膨胀等同于沿边界扫过一个圆。因为边界的不规则性，被圆扫过的条带面积不会直接增大半径。在对数坐标中画出由圆扫过的面积（有时也称为香肠）与半径的关系图，会产生一条直线，由该直线的斜率可得出分形维数。这种闵可夫斯基技术比理查得森方法的历史更悠久。它能很好地处理像素图像，使用欧氏距离图来执行膨胀时，效果会更好。使用特征轮廓任意一侧区域的欧氏距离图的累积直方图来构造对数坐标图时，对特征的方向相当不敏感，且速度远快于迭代膨胀，因为它会根据像素离边界的距离来对该像素赋一个灰度值。这种方法的例子已在第 8 章中给出；图 11.15 显示了同一科赫曲线的测量。



(a)



EDM值	像素数	累加	对数(R)	对数(A)
1	9330	9330	0.00000	3.96988
2	5956	15286	0.30103	4.18429
3	5198	20484	0.47712	4.31141
4	4052	24536	0.60206	4.38980
5	4448	28984	0.69897	4.46216
6	4008	32992	0.77815	4.51841
7	4070	37062	0.84510	4.56893
8	3416	40478	0.90309	4.60722
9	3580	44058	0.95424	4.64402
10	3620	47678	1.00000	4.67832
11	3640	51318	1.04139	4.71027
12	3796	55114	1.07918	4.74126
13	3096	58210	1.11394	4.76500
14	3260	61470	1.14613	4.78866
15	3074	64544	1.17609	4.80986
16	3114	67658	1.20412	4.83032
17	2978	70636	1.23045	4.84903
18	3082	73718	1.25527	4.86757
19	3072	76790	1.27875	4.88530
20	3000	79790	1.30103	4.90195
分形维数=2－斜率=			1.28257	

(b)

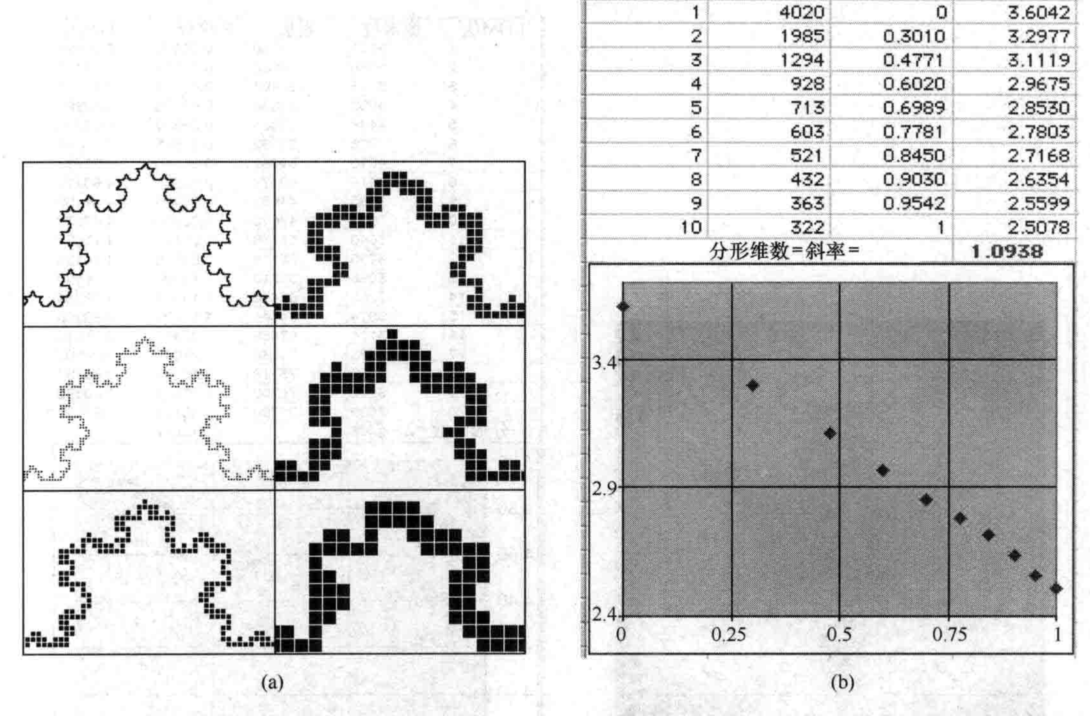
图 11.15 (a)围绕科赫曲线轮廓的欧氏距离图的累积直方图；(b)闵可夫斯基测量产生的分形维数

闵可夫斯基方法产生的维数不一定与理查得森方法（或其他分形维数过程）产生的维数相同，所以在比较由不同标本得到的维数时，仅使用一种方法很重要。由于像素图像的有限分辨率，对于科赫曲线分形，这两种方法都不能给出理论上期望的准确结果，但两者的结果相当接近。

分形维数测量的第三种方法产生的另一个维数，与其他方法所得值的意义相同，但数值上有较大差别。柯尔莫戈洛夫维数是通过对网格计数得到的。在图像上绘制线网并统计边界穿过的网格单元的数量。在对数坐标系中画出这个数字与网格大小的关系图时，由直线的斜率可再次得到一个分形维数，如图 11.16 所示。通过逐步将图像加粗为 2×2、3×3 等的像素块，可自动计算出像素图像的维数，这有时称为镶嵌合并。这也许是最快的方法，但准确度和数值精度最差，因为它的步数最小，进而导致生成直线的点最少。

但分形维数确实产生了一个数值，这个数值概括了特征边界的不规则性或“粗糙度”。这是人们注意到的特征“形状”的一个方面，且是人们用于分类或比较特征的一种定量方法。图 11.17、图 11.18 和图 11.19 给出了几个自然分形的例子。对淀积涂层（Sander，1986；Vicsek，1992）、各种化学或物理腐蚀

形成的表面（Sapoval，1991）、某些类型的断裂和磨损（Ray & Mandal，1992；Przerada & Bochenek，1990；Mecholsky & Passoja，1989；Wehbi et al.，1992）等已报告了分形维数。分形形状也可用于癌症检测诊断（Bauer & Mackenzie，1995；Baish & Jain，2000；Kikuchi et al.，2002；Jayalaitha & Uthayakumar，2007；Martin-Landrove et al.，2007；Abdaheer & Khan，2009；Uppal et al.，2010）。在更大的尺度下，一些湖泊的形状是分形的，而且维度被观察到与高度相关（Goodchild，1980）。如前所述，云是分形的（Lovejoy，1982），海冰图案也是分形的（Ivanov et al.，2007），河流图案、树枝、肺部的气管等也是分形的。



单元格的宽度	计数	对数(W)	对数(N)
1	4020	0	3.6042
2	1985	0.3010	3.2977
3	1294	0.4771	3.1119
4	928	0.6020	2.9675
5	713	0.6989	2.8530
6	603	0.7781	2.7803
7	521	0.8450	2.7168
8	432	0.9030	2.6354
9	363	0.9542	2.5599
10	322	1	2.5078
分形维数=斜率=			1.0938

图 11.16　(a)科赫曲线边界穿过的尺寸增大的方框数量统计；(b)对数图。对数图是一条直线，但其斜率未给出形状的准确分形维数

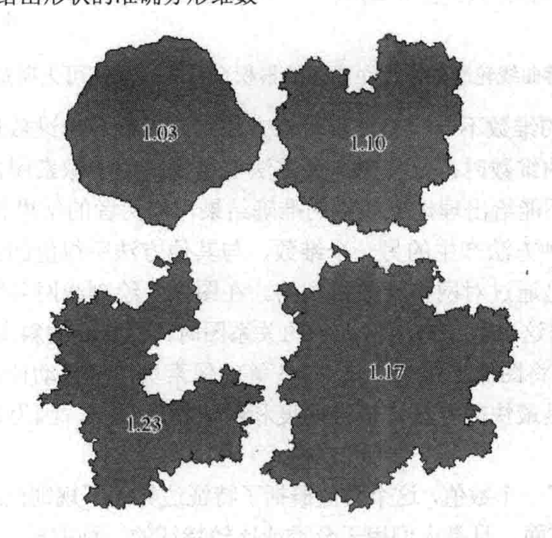


图 11.17　在不同高度采样的尘埃颗粒的分形形状

图 11.18 中显示的截面边界与其三维表示间的关系并不简单。对于截面通过各向同性表面的情况，边界的分形维数比表面的分形维数恰好小 1.0，即截面的拓扑维数与物体体积的拓扑维数之差。然而，表面通常并不是完全各向同性的，且在出现首选方向时，这种简单的关系将不再成立。更为严重的是，如果待测量边界是一个颗粒的投影轮廓，边界的不规则性将会被其他的表面突起部分地遮盖，进而使得测出的分形维数要小某个量级，具体小多少则取决于颗粒的大小和其粗糙度。对此没有常用的校正方法，并且尽管这个程序很常用，但得到的数据（及其后续解译）仍然非常值得讨论。第 15 章将讨论分形表面测量的意义和方法，表面的分形维数通常在 2.0 至 2.999 之间（二维轮廓的维数在 1.0 至 1.999 之间）。

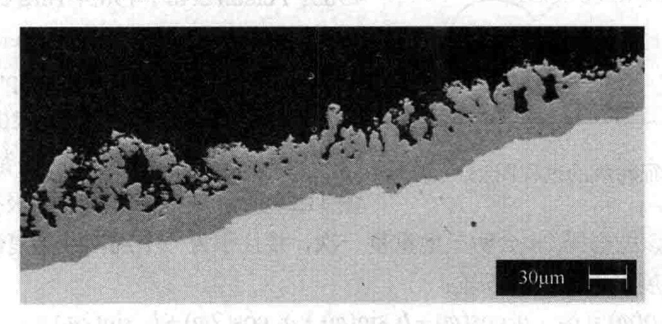


图 11.18　金属上氧化物涂层的截面（分形维数为 1.505）

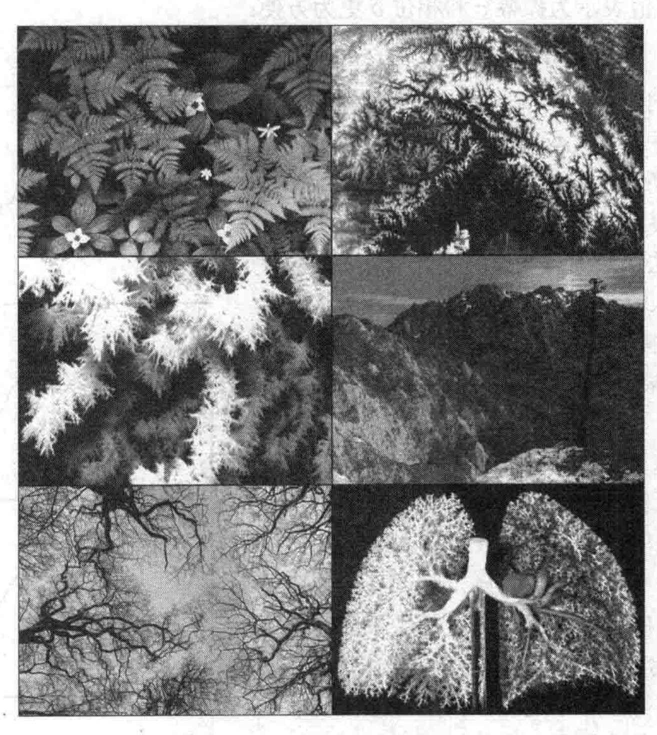


图 11.19　自然分形（从左上角起）：蕨叶和河流排水模式，树突状冰晶的形状，山脉和湖泊，树枝和肺中的气管

11.4　谐波分析

分形维数试图将边界形状的所有细节浓缩为一个以某种特殊方法来描述粗糙度的数字。当然，具有相同的分形维数或局部粗糙度的边界形状有无数种。但另一方面，使用一些数字来详细地保留所有边界信息，可以有效地重建边界的外观细节。

谐波分析也称为谱分析、傅里叶描述子或形状展开，它是数字上描述形状的一种强大方法（Schwartz & Shane，1969；Ehrlich & Weinberg，1970；Zahn & Roskies，1972；Granlund，1972；Beddow et al.，1977；Persoon，1977；Flook，1982；Kuhl & Giardina，1982；Kaye et al.，1983；Rohlf & Archie，1984；Barth & Sun，1985；Ferson et al.，1985；Bird et al.，1986；Diaz et al.，1989，1990；Rohlf，1990；Verscheulen et al.，1993；Lestrel，1997；Haines & Crampton，2000；Bowman et al.，2000；Pincus & Theriot，2007）。经典方法通过将边界转换为极径（角度）的函数形式或 $\rho(\phi)$ 开始。如图 11.20 所示，将特征质心的极径画成角度的函数，并画出长度来展开形状。该图每隔 2π 会明显地重复一次，并且作为一个周期或重复函数直接进行傅里叶分析。这就允许求出级数展开中的 a 和 b 项：

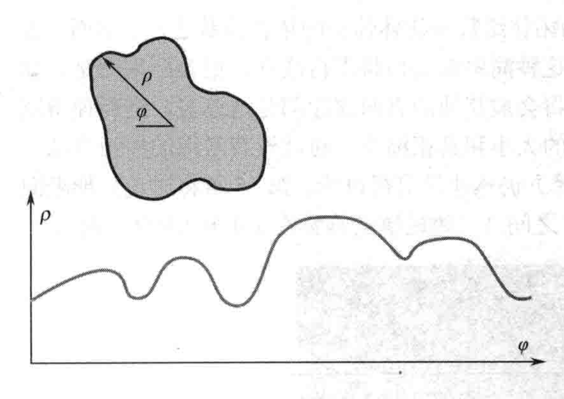


图 11.20　展开特征轮廓的过程图示

$$\rho(\varphi) = a_0 + a_1\cos(\varphi) + b_1\sin(\varphi) + a_2\cos(2\varphi) + b_2\sin(2\varphi) + \cdots \tag{11.1}$$

对每个正弦项，将值表示为振幅 c 和相位 δ 更为方便：

$$c_i = \sqrt{a_i^2 + b_i^2}, \quad \rho(\varphi) = \sum c_i \sin(2\pi i\varphi - \delta_i) \tag{11.2}$$

该级数继续到第 n 项等于沿圆周一半的点数。但傅里叶分析的一个特点是，级数中的前几项即可保留大部分特征形状的细节。如图 11.21 所示，只使用级数的前十多项或二十多项，就可将原始形状画得与原始像素表示的精度一样好。在许多情形下，舍去级数中每一项的相位信息 δ_i 而对每个频率使用单个系数 c_i，对形状描述并不会造成太大影响（但重建原始形状时需要相位信息）。因此，在展开边界的谐波或傅里叶级数展开中，前几个 c 值就包含了关于特征形状的大量信息。

对画出极径与角度关系图的一种反对意见是，非常不规则的部分剖面曲线相对于平滑区域是欠采样的。更自然的一种方法是，画出极径与沿周长的距离关系图。形状展开还存在另一个更为严重的问题。如图 11.22 所示，如果轮廓是凹角或砍口，极径与角度的关系曲线就会是多值的。在每个角度总是使用最小（或最大）极径可消除多个值，但会使得图形不连续。傅里叶分析要求函数是连续的。为了避免该问题，形状展开的另一种方法是，将直线的斜率改画成沿该直线的距离的函数。此图也是一个周期函数，故可用完全相同的方式进行分析，但在某些点处斜率可能会变得无穷大。将斜率的变化画成位置的函数可克服此问题。

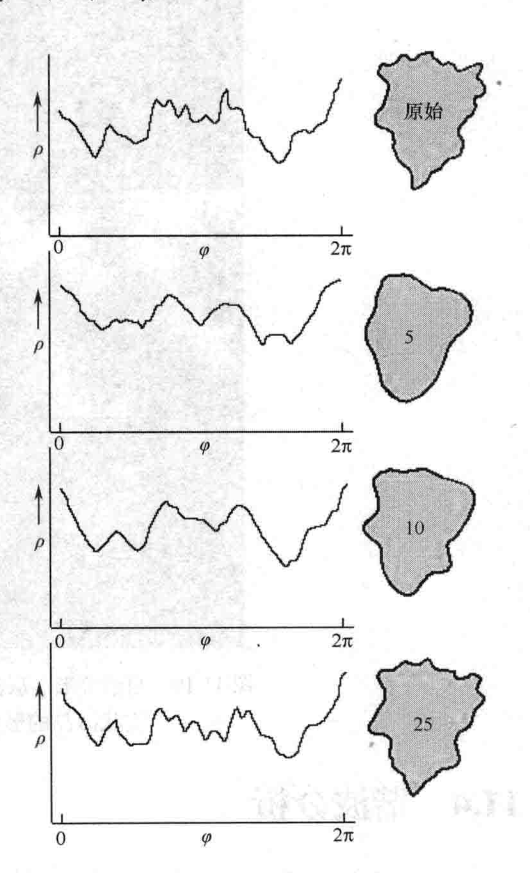


图 11.21　从傅里叶级数展开的前 5 项、10 项和 25 项重建真实物体（一粒宇宙尘埃）的轮廓（Kaye，1983）

特征的链码边界表示已经包含了斜率与沿边界的位置信息的关系，但具有数据非均匀间隔的问题（对角线链长于水平链和垂直链）。解决这个问题的一种有

效方法是，使用 5 个具有同一方向的较短子链来代替每个水平链，并用 7 个这样的子链来代替每个对角链。对于实际用途，比率 7:5 足够接近 $\sqrt{2}$，而沿边界的子链总数仍然相当短。链值（它指明方向）与位置关系图给出了展开的特征形状。对特征沿链码的值序列进行傅里叶分析，可产生包含形状信息的一系列 c_i 值。很难解释这些值，因为方向数不连续。

链码不需要记录为一组对应于 8 个像素邻域方向的值。由方向 8 和方向 1 之间的不连续性造成的分析困难，可通过转换到如图 11.23 所示的差分链码来克服。这只是使用模算术的值间的差，因此零表示在各个已有方向上一直向前处理边缘，正值表示在某个方向增加弯曲角度，负值表示相反方向。采用这种形式，形状是旋转不变的，除非有限的像素网格引起了锯齿，如图 11.24 所示。差分链码通过将角点的绝对值变化在某个距离范围内定义为最小，也可更容易地找到角点（Freeman & Davis，1977）。

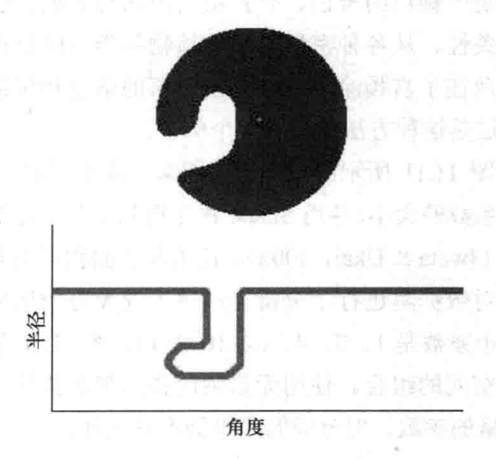


图 11.22　凹角形状及其产生的极径与角度关系的多值图

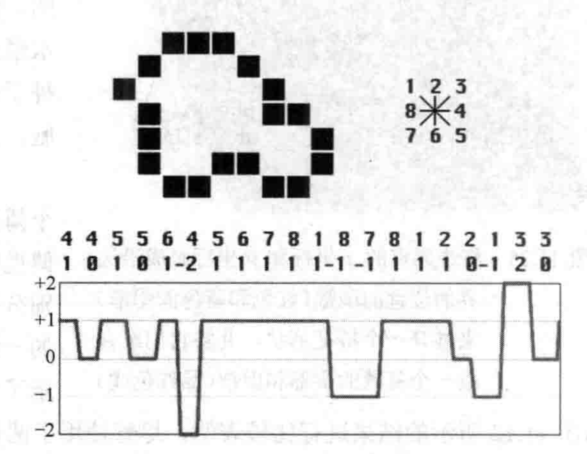


图 11.23　特征轮廓的链码表示，它从蓝色像素开始并以顺时针方向行进，显示了传统的方向码和差分链码，并显示了插入子链的图形，如文中所述

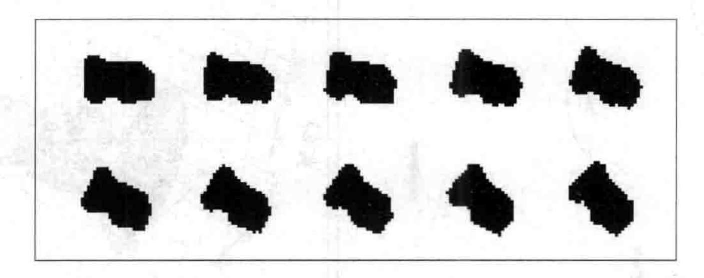


图 11.24　在像素网格上以 5° 步长旋转时，差分链码所定义的一个小形状的细节，显示了锯齿像素

展开的差分链码的数字集可用于计算形状的离散傅里叶变换，但人们在实际中更常使用快速傅里叶变换（FFT，它要快几百倍）。经典形式 FFT（Cooley & Tukey，1965；Bracewell，1989）的工作原理是一种分治策略，它要求原始函数或图中的值的数量是 2 的整数次幂（64，128，256，512，…）。展开的差分链码不可能有满足这一要求的数量，且在任何情形下不同的物体都会有不同的长度。解决办法是沿图内插出固定数量的值，得到具有足够分辨率来表征形状重要细节且长度总是 2 的整数次幂的采样值。

以下例子中使用的一种最普通的谐波分析技术，将沿边界出现的点的 x 坐标和 y 坐标组合为形如 $(x+iy)$ 的复数，并直接在坐标序列上进行傅里叶变换。这可能视为构造特征轮廓的 x 投影和 y 投影

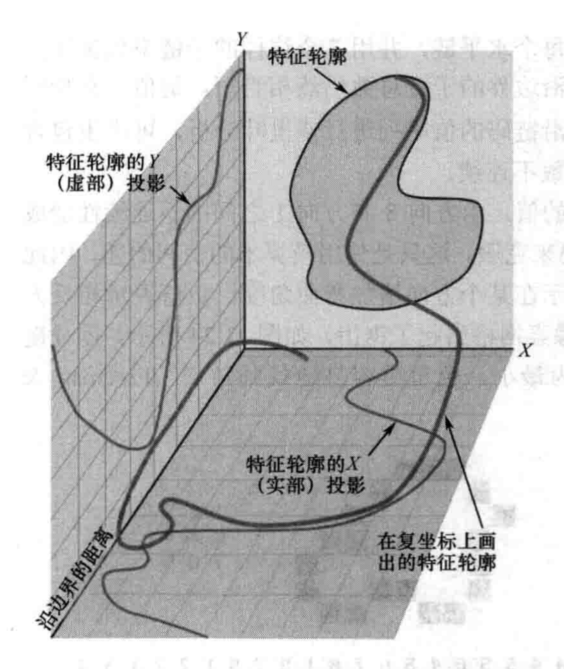


图 11.25　将边界点的 x 坐标和 y 坐标画成沿边界的位置的函数（红色和蓝色的图形）来展开一个特征形状，并将它们组合成一个复数的实部和虚部（品红色线）

图，并将所得值傅里叶变换为沿边界位置的函数，如图 11.25 所示。至于上面的链码变换，这可能需要沿边界内插出 2 的整数次幂个点来使用经典的 FFT，或者使用不需要 2 的整数次幂个点的快速傅里叶变换（Frigo & Johnson，2005；Johnson & Frigo，2007）。结果是概括特征形状的一组 c_i 值。

使用标准的统计测试（如逐步回归法、主成分分析法）可比较物体的类别，以确定哪些项可用于特征分类或识别。在大量情形下，这种方法被证明是成功的。沉积物中颗粒的辨识、有孔虫的形状与它们生长水温的相关性、从各种密切相关的植物种类中区分出种子，以及在子宫颈涂片中分辨出健康的细胞和癌细胞，只不过是这种方法成功的几个例子。

使用图 11.11 所示的相同树叶图像，求出了前 30 个傅里叶系数的大小，并用 Shape 程序得到了归一化的傅里叶项（Iwata & Ukai，2002）。使用与上面相同的判别分析法对数据集进行了分析。产生无误差分类所需的一组最小系数是 1、2、4、6、10 和 14。图 11.26 显示了参数空间的组合。使用无量纲比作为形状描述子与图 11.12 所示的结果进行比较表明，尽管使用了两倍数量的参数，但分类的结果仍不是太好。

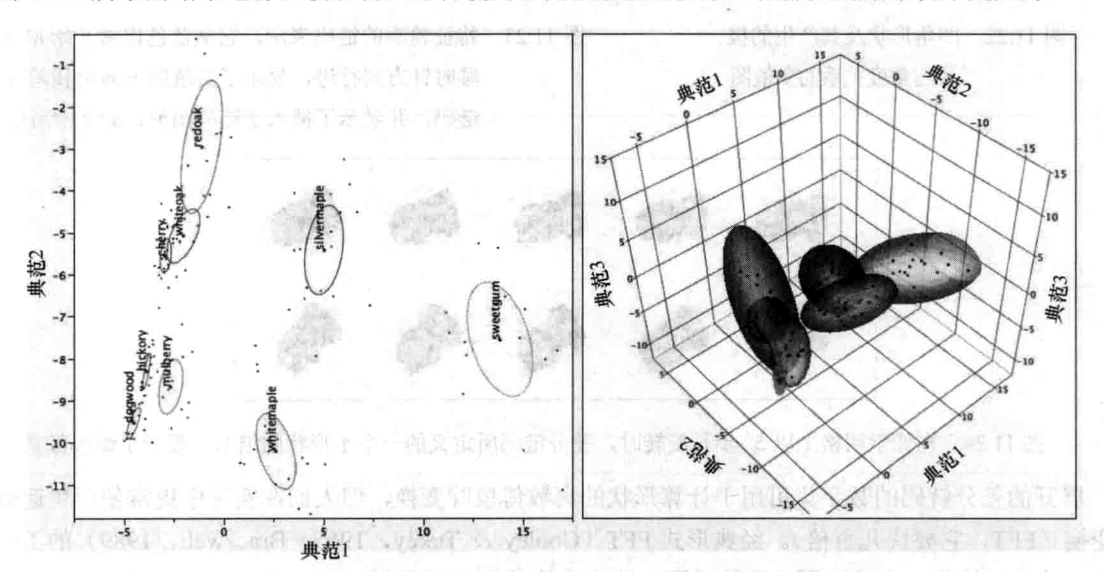


图 11.26　使用 6 个傅里叶系数判别分析叶片样本的结果

为了说明一个更复杂的分类问题，我们考虑经常用到谐波分析的一个领域，即以 1.32 微米/像素分辨率记录的包含近 900 个颗粒的十几个砂粒样本的光学显微镜图像。这些砂粒选自我在世界各地的收藏，如表 11.3 所示。颗粒随机散落在载玻片上，并用透射光来观察剖面。颗粒形状多种多样，完全无法通过原产地来识别它们。与易被观察者识别的树叶不同，砂粒只有很少几个人类可视觉提取并用于分类的特征。图 11.27 显示了几个砂粒样本的颗粒；获取图像后，为了方便展示和测量，每

个二值剖面图已被手工排列在一幅图像上。目标不是识别各个颗粒，而是要通过测量少量颗粒来进行分类。

表 11.3 砂粒样本

沙特阿拉伯鲁卜哈利（空地红砂）	北卡罗来纳州海洋岛（日落海滩南）
北卡罗来纳州 Brevard 毗斯迦山国家森林公园（戴维森河）	死海谷（以色列 Moshav Hatza）
北卡罗来纳州 Wrightsville 海滩	夏威夷（火山国家公园，Waha'ula 游客中心，Puna 海岸，大岛）
犹他州锡安国家公园（West Wall Kayenta Formation）	大巴哈马岛（普罗维登斯海峡西北）
科德角（马萨诸塞州，Wellspoint，马可尼站）	冰岛（Dyrholoey 东海岸海滩）
新墨西哥州白砂	毛里塔尼亚（Akchar Eng，残遗砂丘）

砂粒不是矿物，其颗粒尺寸范围为 0.05～2mm（较细的颗粒称为粉砂，较大的颗粒称为砂砾或卵石）。矿物由多种砂石组成，并不限于海滩上的石英（海滩上的砂粒也不仅仅是石英）。在我的收藏中，砂粒的种类有石英、碳酸钙、石膏、橄榄石、玄武岩和其他矿物质。纳米比亚有一个著名的钻石海滩。许多砂粒中包含了多种矿物质颗粒。

砂粒在海滩上随着波浪作用反复来回地滚动并被该过程分级，因此它们是大小和硬度基本一致的颗粒，这些颗粒完全不同于河床中形成的散乱颗粒。砂粒经风蚀后，表面上形成了许多凹陷。新砂有尖锐的棱角，完全不同于圆滑的旧砂。

地质学家使用与自然风化相关的棱角和圆度来描述砂粒的形状特性。图 11.28 给出了由肯特州立大学 Σ-Γ-E（Sigma Gamma Epsilon）联谊会编制的部分砂粒分类指南。注意，这些术语并不基于测量，而基于砂样或形状的视觉比较。

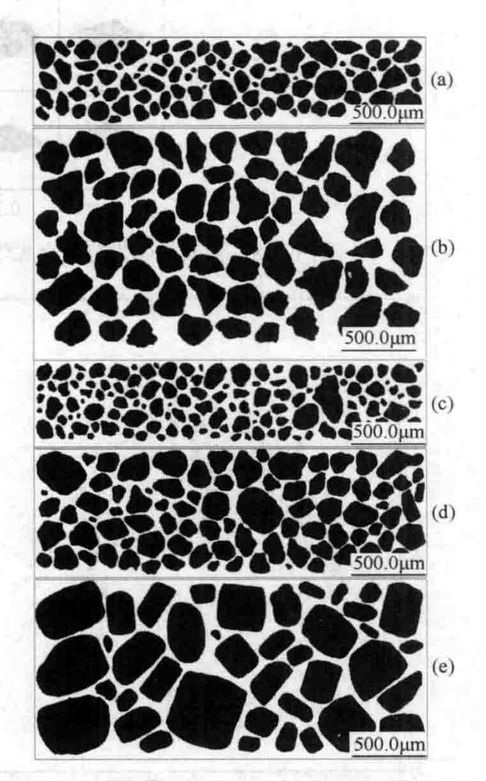


图 11.27 砂粒的二值图像：(a)沙特阿拉伯鲁卜哈利的红砂；(b)北卡罗来纳州 Wrightsville 海滩的砂粒；(c)犹他州锡安国家公园的砂粒；(d)毛里塔尼亚 Akchar Eng 的砂粒；(e)新墨西哥州白砂

对表 11.3 中列出的砂粒二值图像进行测量，得到了一组经典的形状描述子（表 11.1 中的无量纲比），以及第 10 章中描述的各种测量尺寸。将这些描述子和测量尺寸组合到一个数据集中，可进行判别分析，就如同对树叶进行的分析那样。使用凸面积、凸周长（两个尺寸测量）和实心度（面积与凸面积之比）逐步加入和去除各项，得到了最佳的结果。如图 11.29 中的图形所示，这种方法并未完全分开所有类型的砂粒，而是约有 66%的砂粒被错误地分类。

但使用这种方法仍然可以区分每对样品的不同。对少量有代表性的颗粒进行测量，可明显地将它们分为几组（见图 11.30），甚至成功地对多数颗粒进行了分类。例如，使用周长、凸周长和实心度，正确地分类了 90%的沙特阿拉伯鲁卜哈利的红砂和北卡罗来纳州海洋岛的海砂。使用凸面积、等圆直径和分形维数，百分之百地分类了夏威夷砂粒（新砂）和毛里塔尼亚砂粒（老砂）。使用圆度、实心度和半径比，正确地分类了 87%的新墨西哥州白砂和毗斯迦山河砂。对每对样本依次进行这种分类，识别出每种情形下的不同测量参数，进而使用这些测量参数进行分组。

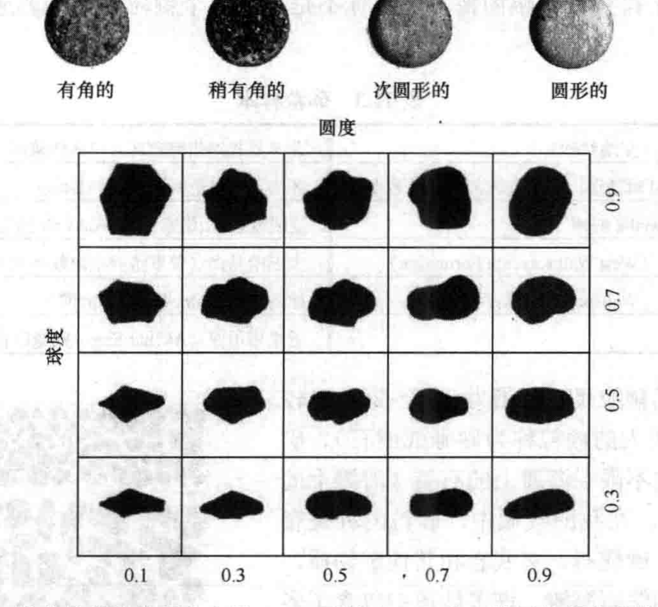


图 11.28　俄亥俄州肯特州立大学 Σ-Γ-E 联谊会 Γ-Z 分会编写的部分砂粒分类指南

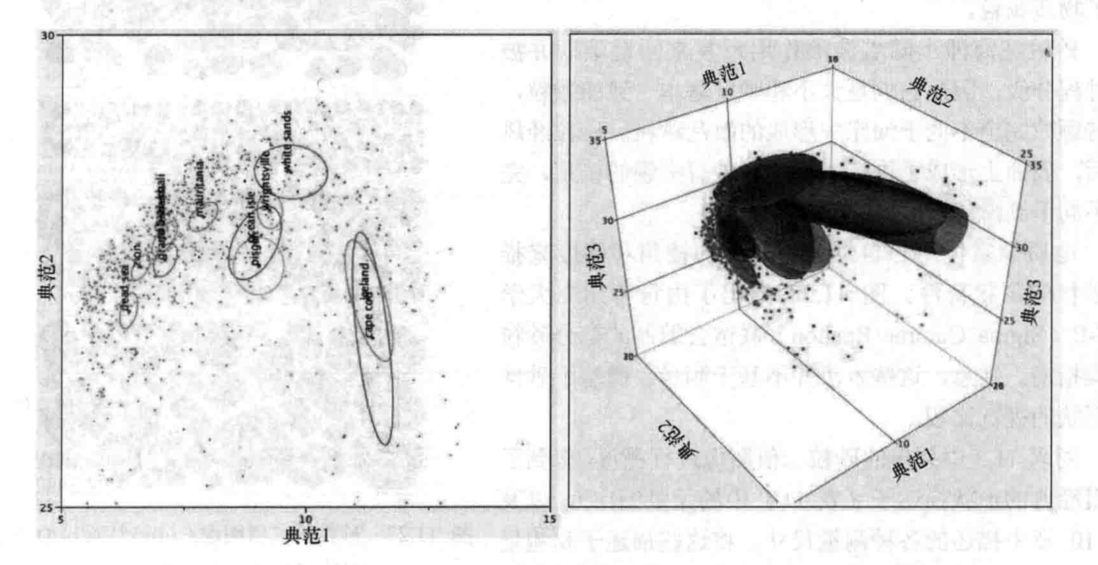


图 11.29　使用三种经典的尺寸和形状测量法对砂样进行判别分析的结果

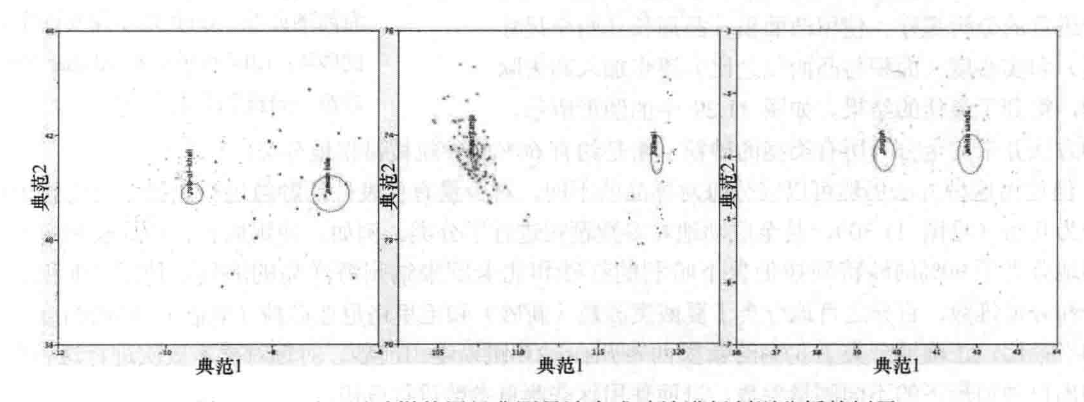


图 11.30　对一对砂样使用经典测量法来成功地进行判别分析的例子

使用傅里叶形状描述子也可进行相同类型的分析。对沉积物颗粒应用傅里叶形状描述子已有很长的历史；在某些情形下，使用傅里叶描述子对不同河流中的侵蚀颗粒成图，可给出沉积物沉淀的位置。就像分析树叶的过程那样，对每个颗粒包含了傅里叶级数中前 20 个幅度项的数据集。如图 11.31 所示，用第 3 个、第 4 个、第 11 个和第 15 个级数项得到了最好的结果，但并未分开各组（78%的颗粒被错误地分类）。增加更多的级数项并未改进该结果。

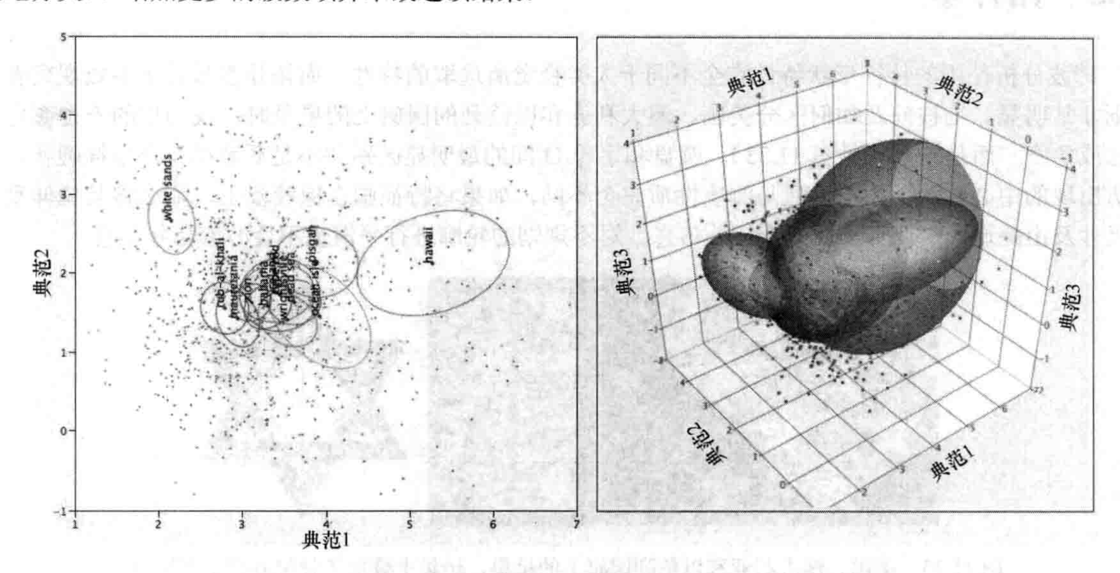


图 11.31　使用 4 个傅里叶系数对砂样进行判别分析的结果

成对考虑两组砂样时，可区分两组砂样（见图 11.32）。使用几十颗砂粒，清楚地分开了两组砂粒，且多数颗粒被正确地分类。该图显示了上面给出的同一对砂样的结果。使用系数 1、5、10、11 和 14，正确地分类了 17.9%的鲁卜哈利砂粒和欧逊岛砂粒；使用系数 1、8 和 13，正确地分类了 15.9%的夏威夷砂粒和毛里塔尼亚砂粒；使用系数 8、9 和 12，正确地分类了白砂和毗斯迦山砂粒。对于经典的无量纲比测量情形，由于每个成对的比较可使用不同的组合项，这就使得综合解释结果非常困难。

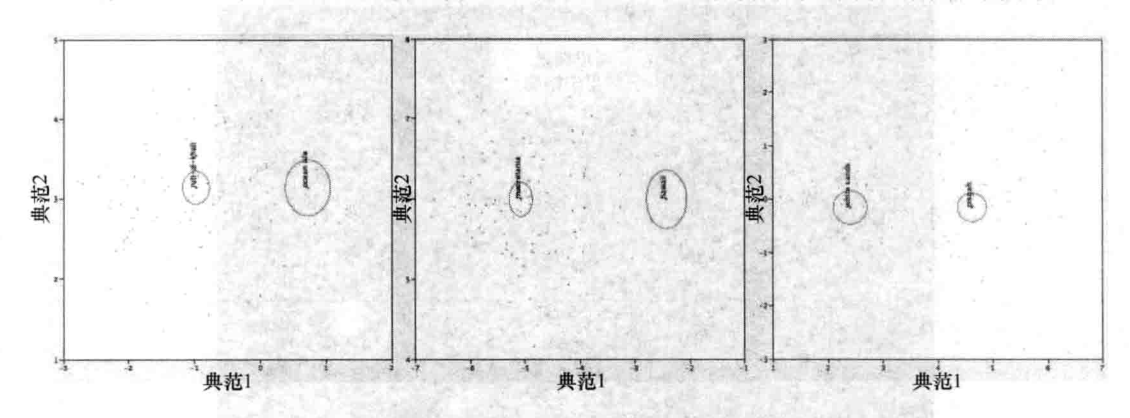


图 11.32　使用傅里叶系数对一对砂样成功进行判别分析的例子

尽管谐波分析方法很成功，但多数图像分析人员并不将它作为常用工具，多数图像处理与测量软件中也未包含它。造成这种被动局面的部分原因是，谐波分析需要非常大的计算量才能求出参数，而对这些参数的解译需要大量的统计分析。但随着计算机能力的持续提升和成本下降，这已不再是主要原因。更可能的原因是，人们对频率项不熟悉，因为人类视觉无法关联这些频率项。我们直观地从图像中提取的形状信息并不能揭示这些数值因子。基于第 8 个和第 13 个谐波系数的两种砂粒的区分，一

定程度上可以理解为物体不规则边界中那些频率项的幅度差别，但人类观察者无法理解其中的关联关系。该方法的成功说明了基于计算机的测量算法的能力在定量（准确度和精度）和定性（可以测量的物体）方面都超过了人类，但这并不意味着使用这样的工具时人们会感到舒适。

11.5 拓扑学

谐波分析使用的特征形状特性完全不同于人类视觉所选取的特性。而拓扑参数对于多数观察者而言则非常明显。当被问及如何区分美国、澳大利亚和以色列的国旗上的星星时，我们指的不是颜色、尺寸或角度，而是点数（见图 11.33）。圆盘和字母 O 间的最明显区别并不是后者稍有点呈椭圆状，而是所出现的中心孔洞。拓扑性质与测度性质完全不同。如果将特征画在橡胶板上，那么将其拉伸到任何尺寸及由此造成的变形不会改变拓扑信息。对不规则的轮廓进行平滑也不会改变拓扑信息。

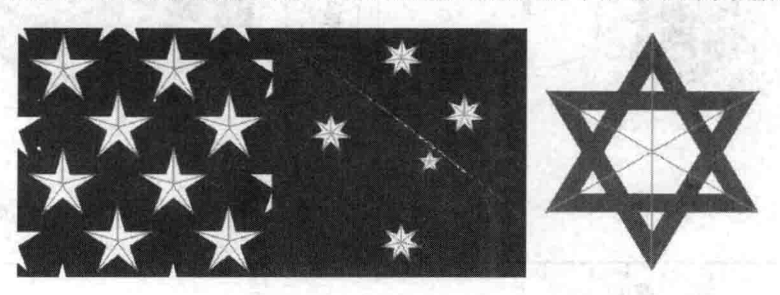
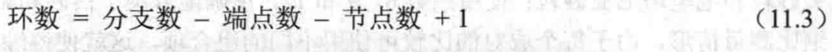


图 11.33　美国、澳大利亚和以色列国旗上的星星，星星上叠加了骨架并标记了端点

拓扑学起源于 1735 年欧拉对历史上著名的“哥尼斯堡七桥”问题无解的证明。如图 11.34 所示，该问题求的是穿行该城的步行路线，要求只经过每座桥一次。欧拉证明这是不可能的，并在求解过程建立了连接节点、分支和端点数量的简单规则，这个规则既适用于二维图形或图像，也适用于三维实心的边缘、顶点和表面。在二维情形下，欧拉法则是



$$\text{环数} = \text{分支数} - \text{端点数} - \text{节点数} + 1 \qquad (11.3)$$



图 11.34　哥尼斯堡（现加里宁格勒）地图，显示了河上连接岛屿和城市其他部分的七座桥（显示为红色）

我们可以在二维情形下使用字母表中的字母进行简单的练习。如图 11.35 所示，几个字母具有相同属性的环、端点和节点，因此一个字母可持续扭曲成另一个字母。对于衬线字体，从骨架中删除较短的外部分支可简化形状。可以非常容易地确定特征的拓扑性质意味着，使用拓扑性质区分物体时，拓扑性质是进行分类的主要选择。

特征的一些拓扑性质可以直接由像素阵列（如内部孔洞的数量）来确定。其他性质，如在图 11.33 中所示星星的数量，使用骨架是最容易确定的。骨架由沿特征中线分布的像素组成。仅与单个相邻像素相接的像素是端点，而具有两个以上相邻像素的像素则是分支或节点。统计分支、端点和环（孔洞）的数量，即可得到特征特性的一个紧凑拓扑表示。第 8 章中给出了使用骨架端点来表征形状的几个例子。

不考虑大小、字体或样式的变化，而只考虑我们识别字母表中印刷字母的能力表明，这些特征的特性对视觉分类和识别非常重要。很多将文本图像转换为存储在文件中的字母的字符识别程序大量使用了拓扑规则。

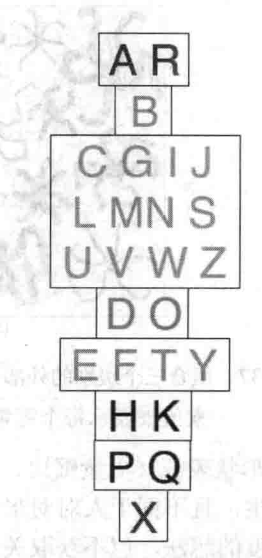


图 11.35　以普通无衬线字体印刷并经拓扑排列的字母表中的字母。每组字母具有相同数量的端点、环和节点

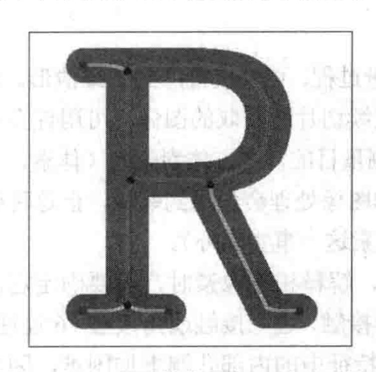


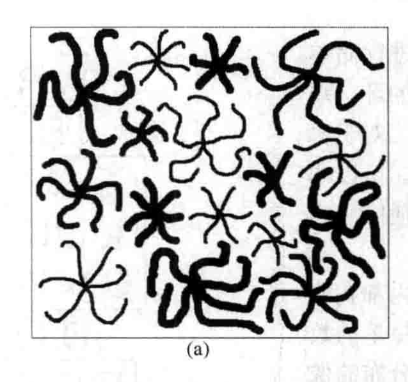
图 11.36　一个形状的骨架，显示了端点（红色）、节点（蓝色）、内部分支（绿色）和外部分支（黄色）。式（11.3）将环数（1）计算为［分支数（8）– 端点数（4）– 节点数（4）+ 1］

图 11.36 说明了可以区分开来并可用于拓扑形状表征的骨架的各个组成部分。有些组成部分，如端点的数量、环或孔洞的数量以及节点的数量，只须通过计数确定。其他组成部分，如外部分支（在端点处终止）和内部分支（从一个节点延伸到另一个节点）的相对长度，则需要测量并用于上面介绍的无量纲比中。组合骨架与欧氏距离图，可以确定不规则结构的宽度。

图 11.37 使用几个特征说明了这个问题，这些特征中弯臂的数量、长度和宽度都是不同的。这些特征十分复杂，直接进行视觉识别和标记很困难，且很难找到三个参数值都相同的一对特征。使用骨架和 EDM 测量每个特征的形状参数可指出哪两个特征是一对。通过将三个测量的参数（臂数、臂长和臂宽）赋给不同颜色通道中的像素值，图中显示了识别结果。具有相同颜色的两个特征，其所有三个属性都匹配。

特征的另一个拓扑性质是其边数。在空间填充结构（比如组织中的细胞、金属中的颗粒或航拍图像中的田地）情形下，每个特征的边数是与该特征相邻的特征的数量。根据相邻特征的数量来标记该特征，可揭示结构的有趣属性，如第 10 章中所示。

使用另一种方法也可求出一个特征的边数。如果将角点定义为边界方向中任意的突变，则可由链码直接统计出边数。对边界像素中的较小不规则性不敏感的一种方法使用凸多边形或外接多边形。如第 10 章所述的那样，这种多边形通常由固定和相对较大数量的边来构建。例如，以 10° 步长旋转轴可形成一个 36 边的多边形。如果特征形状有一个相对较尖的角，那么该多边形的一些顶点会重合或紧密聚在一起。对可合并的两个顶点间的距离设置任意一个限制（通常表示为多边形总周长的某个百分比），可将靠近的顶点收集到一起。当然，对顶点计数等效于对边计数。对多边形的边的最小长度设置一个阈值（通常表示为总周长的某个百分比），来将它视为特征的一条边的表示，是对边数进行统计的一种等效方法。


(a)

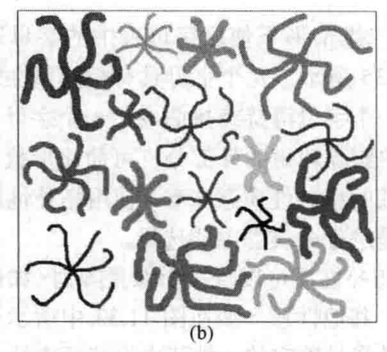

(b)

图 11.37　组合三个级别的外部分支长度、线宽和分支数形成的复杂分支特征。在颜色通道中使用像素值来表示每个测量的参数表明，图(b)中的两个红色特征在所有三个属性方面都匹配

四类形状表征（无量纲比、分形维数、谐波分析和拓扑）都不能令人完全满意。谐波分析对我们来说太陌生，且不同于人对对形状的解译。无量纲比计算方便但不是很具体。分形维数符合人们关于边界粗糙度的想法，但不获取关于形状的更多宏观信息。拓扑正好相反，它可表示特征的总体形状，但不能表示具体细节。在真实情形下，可能需要几个互补的参数。

11.6　三维

上面给出的许多测量过程，以及第 10 章针对二维图像给出的测量过程，可直接推广到三维情形。由共焦光学显微镜、地震成像、医疗或工业断层摄影乃至某些情形下连续切片所获取的图像，可用许多不同的方式来表示，如第 14 章中将要讨论的那样。但对于图像处理和测量目的，立方体素阵列（体素，三维情形下的像素）是最有用的。如同二维情形中那样，这些阵列中的图像处理会使用到邻域，但这种处理所包含的邻域更多，因此需要更多的时间（三维图像中出现更多体素这一事实除外）。

图像测量仍然要求识别出那些彼此相连的像素。在二维情形下，解释相连像素时，需要确定它们之间是 8 连通的还是 4 连通的。在三维情形下，体素接触可是以是面接触、边缘接触或角接触（6 连通、18 连通或 26 连通），并且适用于特征和背景的规则不能相同。识别特征中的内部孔洞更加困难，因为必须测试整个阵列来了解是否存在至外部的任何连接，但逻辑不变。

这种原始的一致性适用于本章和前几章中的大多数测量。对 CT 或 MRI 数据集或一组共焦显微镜图像中的体素的数值求和（或由体素值计算得到的某些属性），很容易就可得到总密度、含水量或已被校准的任何属性。使用体素矩也可测量位置。三维空间中的方向需要两个角度而非一个角度。相邻关系（距离和方向）的意义与解释与二维情形下的相同。

特征面积的三维对应物是特征体积，它可通过统计体素的数量得到。许多方向上的卡尺维度和外接多面体可通过旋转坐标轴并搜索最小点和最大点来确定。当然，在三维情形下获得足够的旋转来将多面体充分地拟合到样本所需的工作量，要比二维情形下的工作量稍多一些。事实上，三维体素阵列中所做的一切都受当前的小型计算机的限制：速度、内存（存储所的体素）、显示（使用体积或表面渲染来显示数据）和人机界面。例如，使用鼠标、轨迹球或其他定位设备，可很容易地选择二维图像中的一个位置。但在三维情形下如何这样做呢？人们尝试过各种方法（有些方法需要实时的立体显示），但没有哪种方法能被人们广泛接受。游戏控制台的发展及熟悉其使用的一代人，或许会在将来给出解决方案。

第 10 章中曾提及二维情形下的周长是一个多少有点麻烦的概念且难以测量。三维情形下对应于特征周长这一概念的是特征的表面积，除了具有周长的所有问题外，它还存在其他一些问题。表面积受

体素分辨率的影响，并随着分辨率的降低而一定程度上加大。如果体素不是立方体，则更难对问题求解。在二维情形下，准确地测量周长比较困难。对于三维表面，边界表示是由三角面形成的多面体的一系列顶点的坐标。由三角形的三个角点计算其面积很容易，但对于任何特定三角形知道使用哪三个点并不容易。作为一个非常简单的例子，我们考虑如图 11.38 中所示的 4 个点。有两种不同的方法在这些点间构建表面，且构建的表面具有不同的表面积（有时会计算出两个表面并取其平均）。

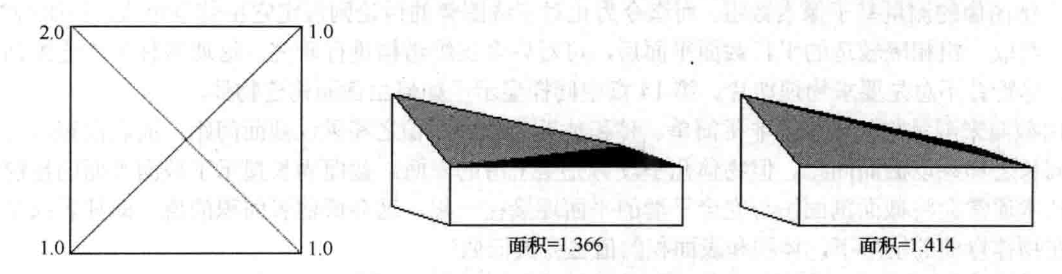


图 11.38　由三角面拼贴成表面的方法有两种。在这个简单的例子中，显示了正方形的 4 个角的高度。将对角线作为公共边，可构建出三角形。如图所示，这将产生一个凸表面或一个凹表面，表面积为 1.366 平方单位或 1.414 平方单位

二维情形下的边界表示要求围绕特征的路径只有一条，而不管该路径多么复杂。而三维情形下却并非如此。边界体素之间的三角面没有必须遵循的唯一顺序。事实上，对于拓扑形状的物体如圆环体（至少有一个开孔穿过它们），并不能保证一个连续表面路径将完全覆盖该表面并到达其上的所有点。这意味着没有类似于链码的对应物，而且基于链码的许多二维尺寸和形状测量，在三维情形下会变得非常困难。

同样，三维骨架也难以得到。对于体素阵列，实际上可以计算两“类”骨架。其一是中线集，其二是中平面集。后者是通过连接每个平面截面的骨架线而形成的，而前者是结点间的连接，或每个平面截面中骨架线的节点间的连接。使用欧氏距离图的三维类似物来找到内切球的球心，通过中轴变换也可得到中线集。任何一类骨架本身都不能获取特征的所有拓扑性质。例如，平面可显示直线骨架上看不到的扭曲和突起。对于三维物体，欧拉数等于物体的数量加上内部腔洞的数量，再减去隧道的数量。

体素线长度的测量与像素线长度的测量类似。由于体素相接的方式有三种，因此须将第 10 章中由式（10.14）给出的规则扩展为

$$\text{长度} = 0.877 \times \text{面接触邻域数} + 1.342 \times \text{边缘接触邻域数} + 1.647 \times \text{角接触邻域数} \quad (11.4)$$

通过在许多方向找到限制点来定义顶点，可以使用大致相同的方式来形成一个外接多面体（类似于二维情形下的外接多边形）。凸状外接多面体的欧拉关系是

$$\text{顶点数} - \text{边缘数} + \text{面数} = 2 \quad (11.5)$$

计算为尺寸比的多数形状参数可简单地推广到三维情形。通过将半径表示为两个角的函数并在二维情形下执行傅里叶展开，也可在三维情形下进行谐波分析。在二维情形下，x 和 y 坐标可以组合成一系列复数，并用在傅里叶变换中。在三维情形下则没有类似的方式来这样做。同样，因为没有链码的类似物或将斜率表示为位置的函数，谐波分析方法会受限于非凹角或其他多个值（在每个角度只有一个半径值）的形状。

分形维数对于处理三维表面和网络非常重要，尽管在多数情形下，在二维截面中测量它们更为容易。事实上，这个一般规律对于多数图像测量任务都成立。处理三维体素阵列实际上非常困难。获得结果所需数组的大小和计算量相当可观。由于内存限制或三维成像技术分辨率的限制，体素分辨率通常要比二维图像的像素分辨率差很多。因此，首选技术是，由二维图像得到所需的信息，然后由体视学将这些信息与三维结构关联起来。这种技术通常更快且更精确。

但这种方法并不适用于所有目的。对于各向异性非常明显的材料，我们需要经过仔细的考虑和权衡，才能在受控方向截取截面，但即使这样，定量描述材料仍非常困难，因此首选方法是三维成像。结构的拓扑性质（如单位体积内物体的数量或网络连通性）在二维截面上根本得不到（尽管第 9 章中描述了使用两个这样的平行平面来对三维拓扑取样的 Disector 技术）。三维拓扑的许多方面只能在三维情形下进行研究，因此需要三维成像、处理和测量。

二维图像的测量基于像素数组，而迄今为止对三维图像的讨论则假定它由体素组成。但如第 14 章所示，获取一组相隔较远的平行截面平面后，可对许多三维结构进行研究。这通常称为“连续切片”技术，尽管并不总是要求物理切片。第 14 章中同样显示了如何由截面构建物体。

由截面来测量物体计算上非常简单。体积估计为截面面积之和乘以截面间距，而表面积可估计为截面周长之和乘以截面间距。但物体几乎没有完全平滑的表面。截面周长揭示了截面平面的粗糙度。重建技术通常会将截面剖面图与完全平滑的平面连接在一起，这会低估表面积的值。即使表面是光滑的，在物体弯曲的情形下，体积和表面积的值也会被低估。

正确对齐截面对于正确理解结构的拓扑以及体积和表面的测量至关重要。第 14 章说明了一些可能的困难。当截面靠得非常时，这些问题会减少。在这一限制下，当截面间距与单个截面图像的分辨率相同时，它等同于一个立方体素阵列。对于测量目标，这通常为首选。

第 12 章　特征识别与分类

识别和分类虽然从一出现就受到了人们的特别关注（Duda & Hart，1973），但它们是“高端”（即需要最复杂算法）图像分析领域的补充。分类涉及建立标准，这些标准可用于识别或区分可能出现在图像中的不同物体。这些标准在形式和复杂程度上有很大的不同，从为每类假设的原型代表的示例图像，到来自测量的数值参数，再到关键特征的语法描述等。第 11 章中给出的树叶分类是使用统计工具和数值的例子。识别是使用这些工具持续找到图像内某个特殊物体或结构的过程。它作用于许多不同的层次，包括在图像中找到一张脸或将该脸与某人进行匹配的不同过程。

计算机在分类方面有时比人类做得更好，因为计算不受非关键参数的随机变化影响，且能提取有意义的统计行为来分离群组。有时，这些群组是有意义的，但在其他情形下，它们可能不是我们所预期的类别。另一方面，人类在识别方面通常要比计算机强很多（或至少更快），因为人类可以察觉到已知的少数几个关键因子，进而识别出熟悉的物体。第 11 章中介绍的树叶人类识别速度很快，且通常不会有差错，几乎可以肯定地说它不依赖于与测量数据的统计分析相同的标准。但对于不熟悉的物体或常见物体的不熟悉视图，人类并不能很好地处理。

各种识别和分类技术有着范围极广的复杂性。在低端，如定位和扫描超市包装上的 UPC（通用产品代码）条码或自动分拣并递送邮件到正确的邮政编码的操作，都基于目标的精心设计，以方便读取并提供检测和纠错。它们通常是通过线性扫描而非本书中采用的获取全部图像来检测的，而 UPC 码是对称的（黑白反转），所以扫描可在任意方向读取。医院的患者识别带上、铁路车厢上及许多其他情形下也会使用条码。同时，二维条码也已被开发出来，它能容纳更多的信息。但读取条码是一个专门的领域，本书并不涉及。

分类背后的基本统计处理与图像分析无关，因为用于分类和识别的参数可来自任何源。关于统计的详细介绍，请参阅 Divijver & Kittler（1980）、Schalkoff（1992）、Haykin（1993）、Stefik（1995）、Hand（1981）、Pao（1989）、Lehman et al.（2005）和其他教材。

12.1　模板匹配和互相关

在二维像素图像中执行识别的最低级功能可能是，对支票应用受限的光学字符识别（OCR）。印在银行支票上的字符限制为 0～9 的数字和一些标点符号。此外，字符都有固定的大小、位置和方向，并设计了一种特殊字体来使得它们易于识别。在这种高度受限的情形下，模板匹配可以最小的计算量快速得到结果。

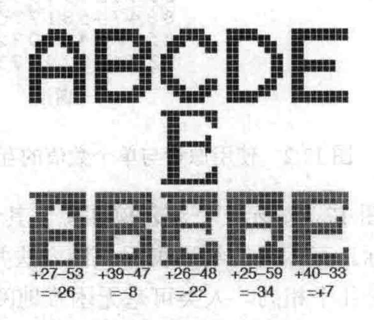


图 12.1　将 5 个字母模板中的每个（顶部）与一个目标字符（中间）匹配，产生了一个净得分（红色匹配像素数减去蓝色不匹配的像素数），正确匹配的净得分最大

图 12.1 中的例子显示了对字母 A～E 进行模板匹配的一个应用。模板由黑色像素组成，这些黑色像素覆盖了存储在内存中的每个目标字符的形状。对待识别的每个字母和存储的所有模板进行“与”和“异或”运算，可统计出匹配的像素数量和不匹配的像素数量。选择给

出最高净得分（匹配数减去不匹配数）的模板来进行识别。在有些实现中，如果净得分对于各个字符变化很大，则将净得分除以模板中的黑色像素数，可使净得分归一化。

在该例中，模板是不精确的，且有足够的宽度和额外像素来覆盖目标字符的适度变化，但显然不能处理大相径庭的字体或样式。只要字符满足内建于模板的假定大小、方向和字体等，几个字母之间形状上的相似性（如 B 和 E）并不会造成严重问题，且这种方法的差错率非常小。出现灰尘等引起的少量噪声时，同样不会出现问题（Secilla & Garcia，1987；Zhang，2009）。

这种类型的模板匹配是一种相当专业且受限的互相关，可用于灰度图像、彩色图像和二值像素阵列。第 6 章将互相关过程介绍为一种技术，该技术通过将目标图像与一幅二次图像匹配来找到亮度模式。如果可建立一组有代表性的目标图像，则该方法可应用于文本图像或任何其他特征的图像。

没有必要使用傅里叶变换来执行互相关。当目标相对较小时，在空间域进行数学运算是有效的。目标模式（旋转 180°）移动到图像中的每个可能位置，并在每个 i, j 位置，根据式（12.1）对互相关值求和来得到一个结果，如图 12.2 所示。

$$\frac{\sum_{x=0}^{s}\sum_{y=0}^{s} B_{x+i,y+j} \cdot T_{x,y}}{\sqrt{\sum_{x=0}^{s}\sum_{y=0}^{s} B_{x+i,y+j}^2 \cdot \sum_{x=0}^{s}\sum_{y=0}^{s} T_{x,y}^2}} \tag{12.1}$$

式中，B 和 T 分别是图像和目标的像素亮度值，求和在目标的整个大小 s 上进行。在目标和图像中明亮像素相似的位置，结果值非常高，而对于目标或图像的整体亮度变化，分母归一化了结果。除了提供匹配度的定量测量外，可适当缩放相关得分来生成一幅导数灰度图像，该图像中的亮度是目标在图像中的每个位置匹配是否良好的一个测度。为检测出现的匹配，该图像可使用顶帽滤波器进行处理，或对该图像进行阈值处理。

```
273988999347777
169777226863788      586653325767 7
884344924776533      764634366376 4
595654879169484      646366572583 3
947178683495148      543587711843 4
555859563987931  357 435666627725 3
251725634726494  595 535256535463 4
836369422783746  753 353591045755 5
455884199783686      667710477635 6
755415676823233      764236656335 4
533566267995831      455743676547 3
596523999583624      767516766654 0
884745817766595      745575157653 6
454864279724597      447561385456 7
737488293735571
     图像         目标      结果
```

图 12.2　使用像素与单个数值的互相关示例。最佳匹配标记为红色（注意这不是一个精确匹配）

图 12.3 显示了一幅示例图像，其中对字母 A 到 E 使用了几种不同的字体。有标记的字母 B 被用做目标。互相关的结果表明，该方法并不能匹配以其他字体出现的相同字母，而对一些其他字母的匹配效果几乎相同。人类可毫无困难地阅读以任何字体出现的字母，即使其拓扑结构有一定的变化。

互相关寻找特征形状和亮度的具体细节，当某个特征被部分遮挡时，尽管互相关是相当鲁棒的，但仍不能容忍大小或方向的变化。图 12.4 显示了相同的目标字母 B，此时应用它来找到具有相同字体的字母，但图像中叠加了噪声图案、线和点。对互相关结果应用顶帽滤波器，并使用彩色比例尺来显示与每个字母匹配的强度。找到了所有的字母 B，与其他字母相比，它具有最大的值。

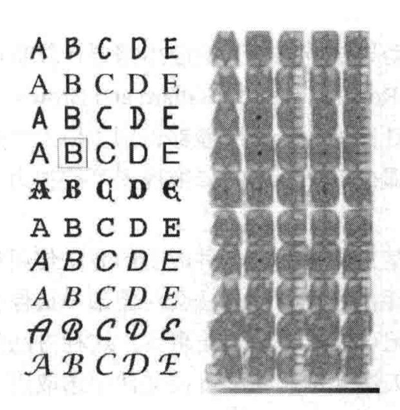


图 12.3　不同字体字母的互相关。加有外框的字母 B 被用做目标，互相关值归一化后的结果值用所示的颜色比例尺编码

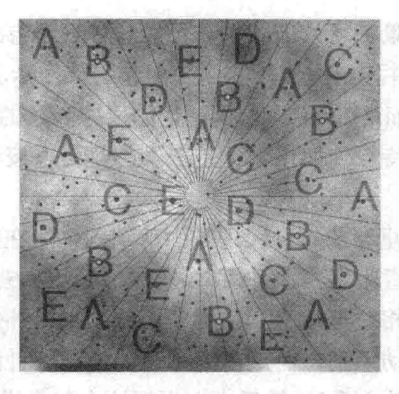


图 12.4　互相关：使用图 12.3 中相同的目标字母 B 找到所有字母 B，而不管可变的对比度、叠加的噪声、线和点。彩色比例尺显示了互相关值的大小

图 12.5 显示了该技术的用途之一——军事目标识别的一个例子。目标图像（一架 F-15 战斗机）产生一个互相关值，该值旋转 10° 下降到 84%，旋转 20° 下降到 80%。两个数值都高于与不同飞机的相关，它对一架 F-14 为 72%，而对一架 B-1 为 47%。在进行计算之前，图像必须按比例缩放为同样大小。在实践中，要么旋转图像使得它们对齐，要么以不同的方向来使用目标图像（如以 15° 间隔旋转）。可能还需要其他方向的视图（侧视图、前视图和斜视图等）。

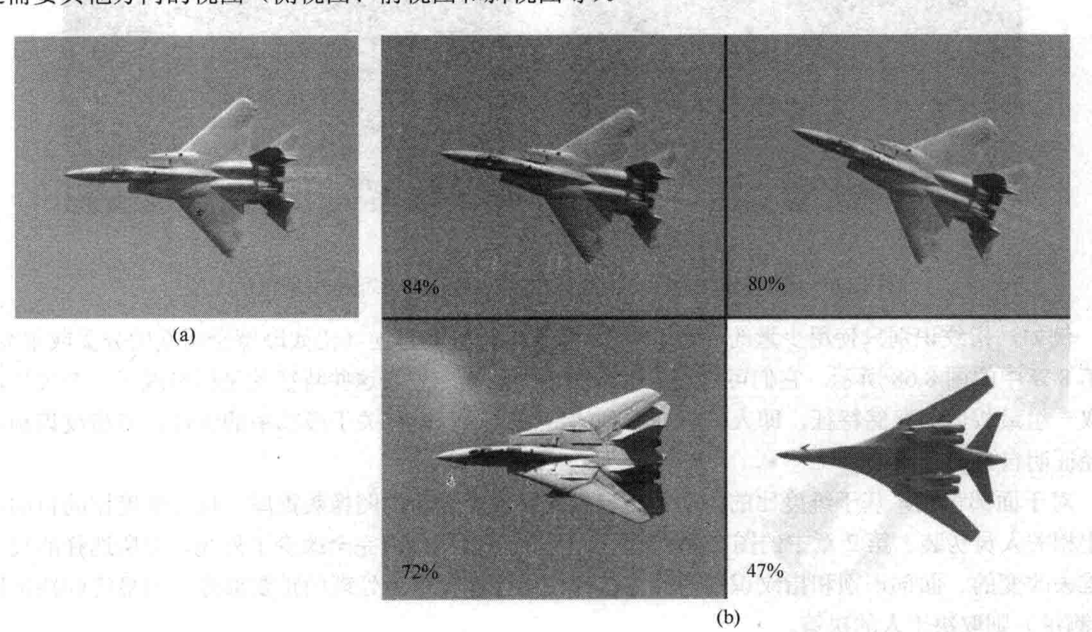


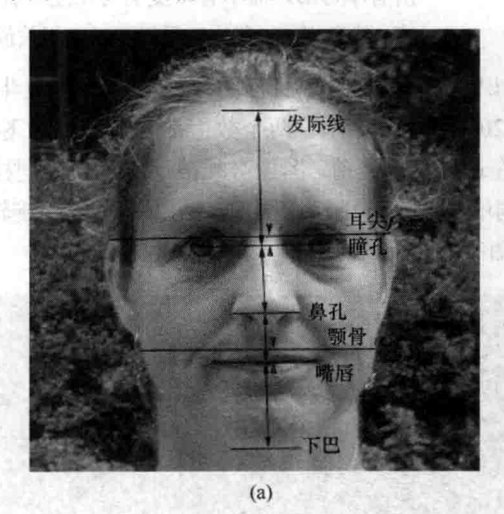
图 12.5　飞机的互相关：(a)目标图像（F-15）；(b)四幅其他图像的匹配得分，显示了形状和方向的影响。顶部两幅图像与目标图像相同，只是旋转角度不同；底部两幅图像是完全不同的飞机类型（F-14 和 B-1）

12.2　参数描述

图像中特征的识别涵盖一个非常宽的应用范围。在某些情形下目标是完全已知的，并且可以由一幅或多幅图像完全表示，所以互相关是一种合适的方法。在其他情形下，目的是让计算机“理解”自

然三维场景，场景中的物体可以各种表示方式出现。诸如自动导航或机器人的应用需要计算机能够提取表面，并按层次连接它们来构建三维目标，进而识别它们（Roberts，1982；Ballard and Brown，1982；Ballard et al.，1984）。这里讨论的主题和目的是非常有限的：让图像分析系统能够识别本质上二维场景中的离散特征。如果物体是三维的且可以按不同方向出现，那么每个不同的二维视图可被视为不同的目标物体。

变得重要的一种应用是面部识别，如在监控视频中找到某人的面孔。一种成功的方法使用选定标记间的垂直距离和水平距离之比，如图 12.6 所示。通过使用距离之比，该方法会对面部与摄像机间的方向变得相对不敏感。使用多个维度可以补偿特殊视图中标记可能被遮挡这一事实。这种方法的目的不是全自动识别，而是在高维空间使用各种比率来创建一个矢量（这些比率可在文件中选取固定但少量最类似的面孔），然后由人根据这个矢量进行比较和匹配。这是在许多其他应用中使用的相同扫描方法，其中的一些应用将在下面描述。



(a)



(b)

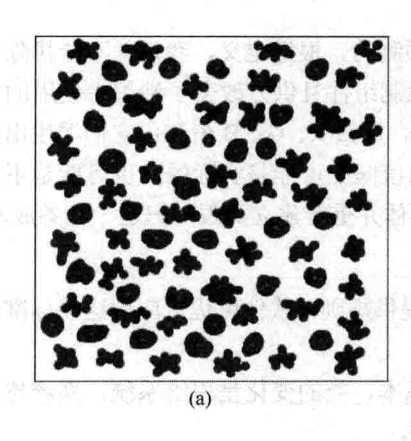
图 12.6　面部识别中使用几个主要的垂直维度和水平维度标记的人脸

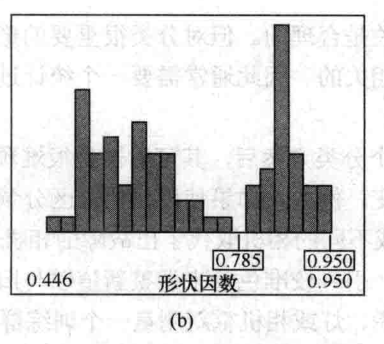
例如，指纹识别只使用少量的"特征点"、细节的位置和方向（比如摩擦脊图案的分叉或末端），如第 8 章中的图 8.68 所示。它们可通过人工或图像处理来定位。这些特征的坐标形成了一个矢量，以选取一组最相似的存储特征，即人们将要看到的特征。尽管没有关于成功率的统计，但指纹识别系统已经证明自身是有用的。

对于面部识别，基于维度比的方法假定已经建立一个合适的图像数据库。选择维度比的目的在于防止相关人员伪装。第 2 章中的图 2.46 显示了尺寸的稍微变化就完全改变了外貌，但所选择的尺寸是很难去改变的。面部识别和指纹识别两种方法都使用计算机作为处理的重要部分，但最终的决策过程（和测量）则取决于人的决策。

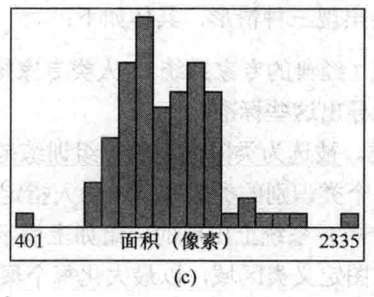
对二维图像的最简特征分类情形（满足大部分实际应用），使用的是前几章中介绍的特定特征测量参数。这些方法称为参数描述。图 12.7 给出了一个简单的例子，其中的特征根据它们的形状可分为两类。这里用到了第 11 章中介绍的几个形状描述子；该例显示了形状因子（$4\pi\cdot$面积/周长^2）。图像中特征的形状因子值的分布显示了两个分离得很好的群体。在两组之间设置界线，可将它们归入各自的类中。图中，"更圆"的特征已用红色标记识别出。

注意，其他测量参数如面积［见图 12.7(c)］未区分出两组特征。并非总能找到一个参数来成功地进行分类，即使找到了这样一个参数时，从许多可能的参数中通过试错来选择最好的一个参数，也可能是非常耗时的过程，但统计分析程序可以为这种处理提供帮助。


(a)



(b)



(c)

图 12.7 (a)一些手绘特征；(b)特征形状因子的直方图；(c)特征面积直方图。形状因子直方图显示了两个分开的类；值大于 0.785 的特征用红点标记，它们看起来比其他特征“更圆”。面积直方图只显示一个分组，因此未区分两组特征

在多数情形下，需要组合几个参数才能成功地进行分类，而使用统计方法可以找到参数的最好组合。作为一个简单的思维实验（Castleman，1996），我们考虑对水果进行分类的一个系统，其中的目标类是苹果、柠檬、葡萄柚和樱桃。苹果与樱桃可以根据其大小（即直径）来区分，柠檬与葡萄柚也可以这样区分。但在某个范围内，苹果与柠檬和葡萄柚的大小相同，因此需要使用第二个参数来区分它们，如区分红色和黄色的平均色调（至少会在排除绿色和金色苹果的情形下有效）。这可显示为参数空间中的一幅二维参数图，如图 12.8 所示。

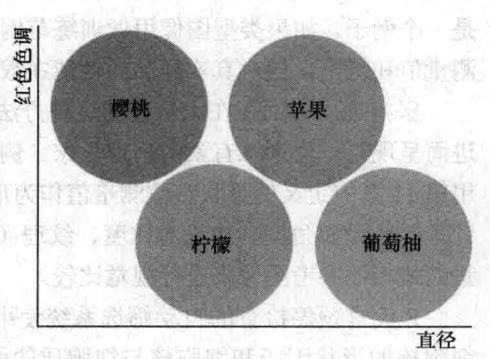


图 12.8 文中描述的水果分类示意图

该例中的各个类是不同的，因此一旦为不同类建立了范围值，就可直接在它们之间画出“决策线”。这通常是通过测量实际样品来进行的。此时，最有效的方法是从每种水果中有意地选取那些极端样本（最大、最小、最红和最黄）来画出边界，而不是仅测量“典型的”样本。

根据测量参数进行决策的目的如下：

（1）能区分各种类，最好根本没有重叠或重叠最少（重叠情形的处理见下方）；

（2）要有最少数量的实用参数，以简化训练过程来确定所有参数组合中的那个期望极值；

（3）参数要彼此无关，以便可由每个参数测量物体的不同特性；

（4）应该是可靠的，以便参数值不随时间变化或由于非受控变量而变化，且在系统硬件变化时要求测量结果一致。

有时，这些目标会彼此冲突。例如，添加更多的参数通常可减少类的重叠，但会使训练过程相当复杂。很难找到彼此真正无关的多个参数。对于苹果，一个参数测量大小，而另一个参数测量颜色，

因此期望它们无关是合理的。但对分类很重要的形状参数可能是彼此相关的，且与由形状参数计算出的大小参数也是相关的。因此通常需要一个统计过程来确定哪组参数在特殊情形下最为合适（例子在下面显示）。

初步建立一个分类方案后，其可靠性是很难预测的。根据定义，我们并不能排除由意外变化导致的物体外观的改变。例如，如果使用颜色来区分预制组件且供方改变了油墨中使用的色素，或荧光灯替代了白炽灯，或不同的相机取代了出故障的相机，那么 R、G、B 值和由它们派生出的任何色调或其他信息也会改变。引入校准色阶来调整新值以与旧值匹配可能是实际的，也可能是不实际的。另一种方法是使用新油墨、灯或相机重新测量一个训练群体并重新建立类限，但这一任务成本通常会很高。

普通分类中会出现三种情形，具体如下：

（1）强制标准（经典的专家系统）。人类专家提供规则。软件可优化应用这些标准的顺序并搜索相关的规则，但不能导出这些标准。

（2）监督分类。被选为类限原型的一组训练范本，类的变化呈现给系统，系统然后开发一种策略来识别它们。为每个类识别的类数和范本由人指定。

（3）非监督分类。系统上出现的一组如上所示的范本，但不告知哪个范本属于哪个类，甚至不告知类数。系统会试图定义类区域，以最大化每个聚类内物体的相似性和群组间的差异。

第二种方法已应用于许多现有工业标准的实现中，这些实现基于专家（或委员会）选择或创建用于指导识别的“类型图像”。对于各个类，它们通常以原型范本的形式（或这些范本的理想草图）出现，但对如何进行匹配未给出特定的指导（且未提供数值测量值或过程）。第 11 章中图 11.28 所示的砂样就是一个例子。如果类型图像用做训练范例（Gomes & Paciornik，2005），那么系统通常可以导出所需的测量值和范围，但对真实样品的性能在很大程度上则取决于如何实际代表那些图像。

医疗应用方面正在大量使用这种方法来自动识别特征，至少作为一种方式来选取可疑或比对图像，进而呈现给一位训练有素的病理学家。例如，为了在乳房 X 线图像上识别肿块，Alto et al.（2005）使用第 11 章中定义的形状参数测量值作为形状因子（4π 面积/周长^2）和凸度（凸周长/总周长），并使用针骨的数量和角度、边缘对比度、纹理（熵）以及密度作为参数空间中的一个矢量，进而选取某个数量的最相似参考图像来进行视觉比较。

P 氏宫颈癌检查的自动筛选系统使用了像细胞核的整体光密度、细胞核的纹理（光密度标准差）、细胞核的形状因子和细胞核与细胞质的面积比这样的参数，来选择载玻片上的少量最可疑细胞，然后呈现给技术人员观察（Luck et al.，1993；Rutenberg et al.，2001）。如果这些细胞中不包含任何表明一种病态状况的细胞，那么载玻片是“正常的”，且无须病理学家进行进一步的检查。但是，尽管这个“自动的”过程通过扫描整个载玻片并选取几个可疑点位来观察提升了速度，但最终的决定仍由人来做出。

还存在其他情形，此时图像的分类不再基于各个特征测量参数的统计，而基于整个场景的特性。最常引用此类辨别的例子，使用了 Brodatz 纹理图像（Brodatz，1966）来表示各种类型的自然和人造表面。图 12.9 中显示了几幅这样的图像。这些图像的傅里叶变换功率谱（见图 12.10）揭示了特性频率和方向，或特定频率处的振幅比。在硬件中使用一组滤波器组选取各个频率并实时进行识别，也可达到同一目的。

这种技术的主要用途似乎是在遥感影像（卫星和航拍照片）的分类中。当与多波段光谱强度比结合使用时，那么如下文所讨论的那样，可使用纹理来识别不同类型的农作物、岩石和其他土地使用模式。但将它们应用到整个场景而非单个特征时，各个数值（如不同频率处的振幅比和方向）将以与本章中所述分类过程中的特定特征测量的相同方式用于决策过程中。

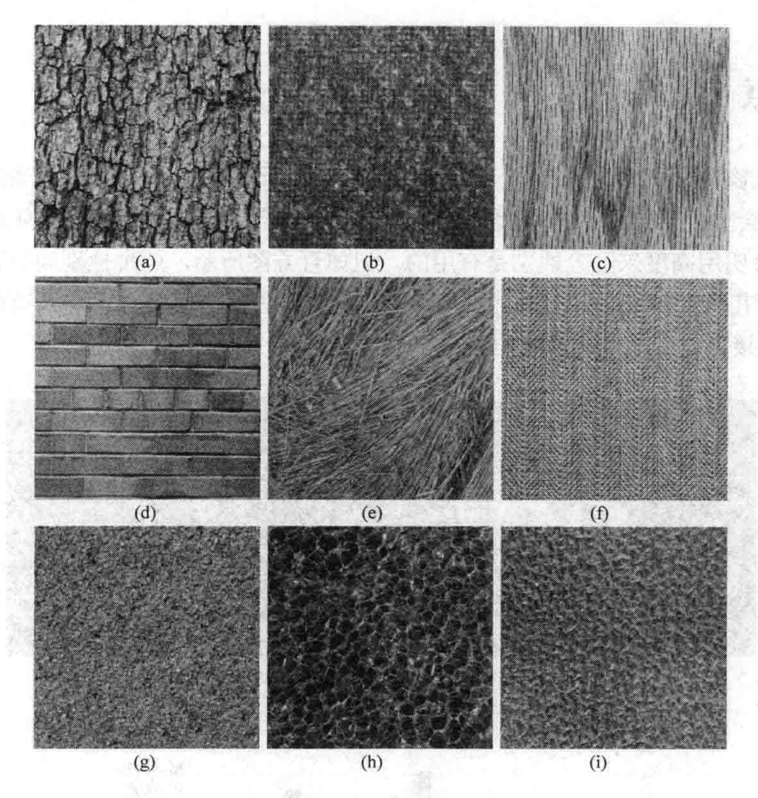


图 12.9　几种 Brodatz 纹理：(a)树皮；(b)毛织布料；(c)木纹；(d)砖块；(e)秸秆；(f)人字形布粒；(g)砂粒；(h)气泡；(i)猪皮

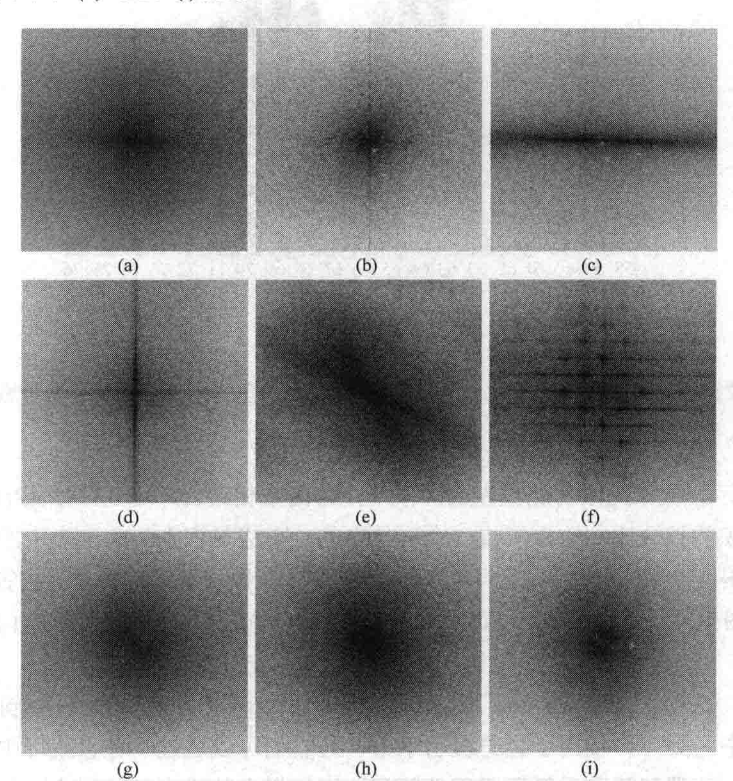


图 12.10　图 12.9 中每幅图像的傅里叶变换功率谱

12.3 决策点

在许多实际情形下，各个类并非完全不同，如图 12.7 中所示的例子。通常，在测量训练群体时，参数值的直方图可能会重叠，如图 12.11 所示。据报道，美国成年男性和女性的平均身高分别为 69.1 英寸和 63.7 英寸，这表明用高度来区分性别是有用的。但如直方图所示，两种性别间有较大的重叠。这种重叠表明需要找到其他参数来更好地进行识别，但避免某种程度的重叠并不总是可能的，此时就需要建立一个能生成可接受误差概率的决策门限。

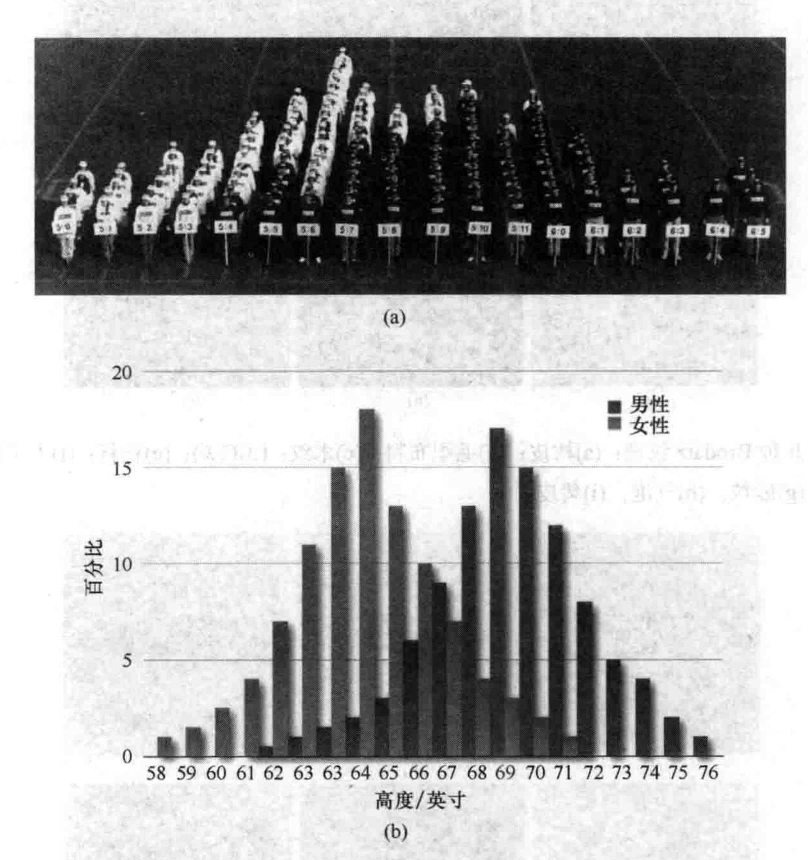


图 12.11　身高数据的直方图：(a)大学生的“活体直方图”（白色表示女孩，黑色表示男孩）；(b)美国成年白人男性和女性的身高

如图 12.12 所示，决策门限与测量参数的概率分布函数相结合，可建立期望的误差率。分布函数尾部的面积（总面积的百分比）给出了将一个群体中的一个成员识别为其他类中的成员的概率。有时决策门限设置为使得两个误差（将 A 错误地识别为 B 或将 B 错误地识别为 A）相等，它不必与最小化总误差相同。对于图 12.11(b)中的身高数据，根据身高进行分类时，66.7 英寸这个门限会错误地分类约 15%的男性和女性。

在其他情形下，必须考虑经济因素。例如，考虑两个大小相同但由不同材料制成的 O 形环。一个环用于拖拉机引擎中，另一个环用于直升机引擎中，且它们可以通过一种查找环中嵌入的彩色标签的成像操作来进行区分。偶尔将直升机所用零件装配到拖拉机上的成本可能很小，此时零件能起作用，

但会造成昂贵零件的浪费。但将拖拉机所用零件装配到直升机上时，则需要承担非常大的责任，因此此类零件会导致发动机失灵甚至更大的故障。此时就需要把将拖拉机零件错误地识别为直升机零件的误差率设置得非常小（一亿分之一或更小），而把将直升机零件错误地识别为拖拉机零的误差率设置得很高（千分之一）。

生物医学成像中也要进行类似的考虑。由于责任问题，人们非常关注自动诊断，例如在P氏宫颈癌涂片中或乳房X光片中使用图像分析来识别癌细胞或病变（Karssemeijer，1998）。广为接受的做法是，将识别可疑特征的判决门限设置得非常低，接受随后可由有经验人员再次进行筛选的很多误报，进而降低疾病的漏报率（Rutenberg & Hall，2001）。

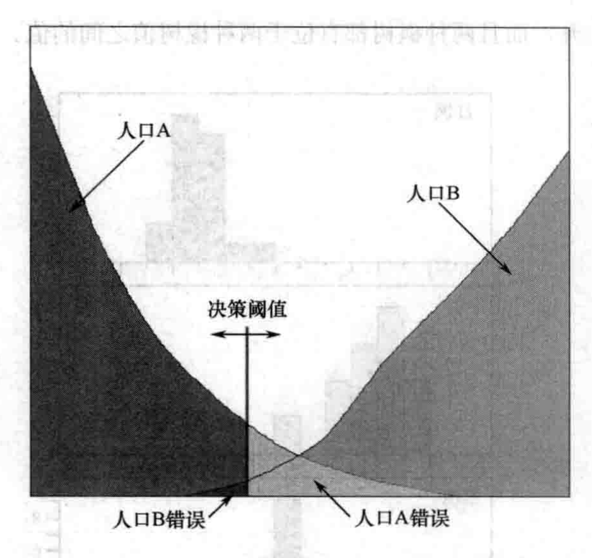


图 12.12　如文中讨论的那样，在分布重叠时设置一个决策点。误分类某个特征的概率，可由每个分布尾部越过决策门限的面积的百分比来测量

12.4　多维分类

在上面给出的例子中，直方图或概率分布函数是单个测量参数的函数。图 12.8 中的图形使用了两个参数（大小和颜色）。此时，通常可以将问题简化为单个导出参数，方法是拟合出一条通过数据的轴线，这通常称为线性判别线或上下文线。如图 12.13 所示，将在训练群体上所测的数据点的分布画在两个参数轴上时，两个不同参数中会有叠加的值，但当投影到另一个派生轴（它是这两个轴的线性组合）上时，则分离得非常好。这条直线可通过使用线性判别分析的回归来拟合。

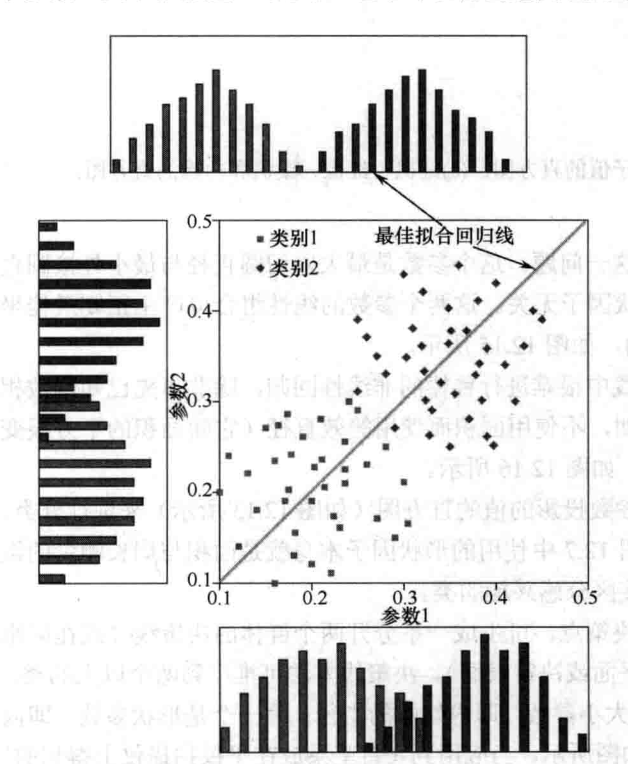


图 12.13　在两个参数轴上均重叠的两个点群，以及一条线性判别线或上下文线，沿该线可最好地分隔两个点群

线性判别分析（LDA）与主成分分析（PCA）有关，第 5 章中将 LDA 描述为一种合并多个彩色通道来得到图像中的最大对比度的方式。PCA 试图找到最大化派生值的方差的测量变量的组合，而 LDA 则试图找到在所测量物体类别间产生最大分离的变量组合。它常用于减少观测参数的数量并找到最有用的组合。

由第 11 章中所示橡树和枫树树叶测量得到的数据可进行详细说明。如图 12.14 所示，只有红枫与红橡树叶的形状因子值的直方图未出现重叠，这表明这两种树可基于单个参数彼此区分开。但在包括了银枫和白橡的数据时，情况就改变了：不仅分布不再分

开，而且两种枫树都有位于两种橡树值之间的值。因此，单个参数无法进行分类。

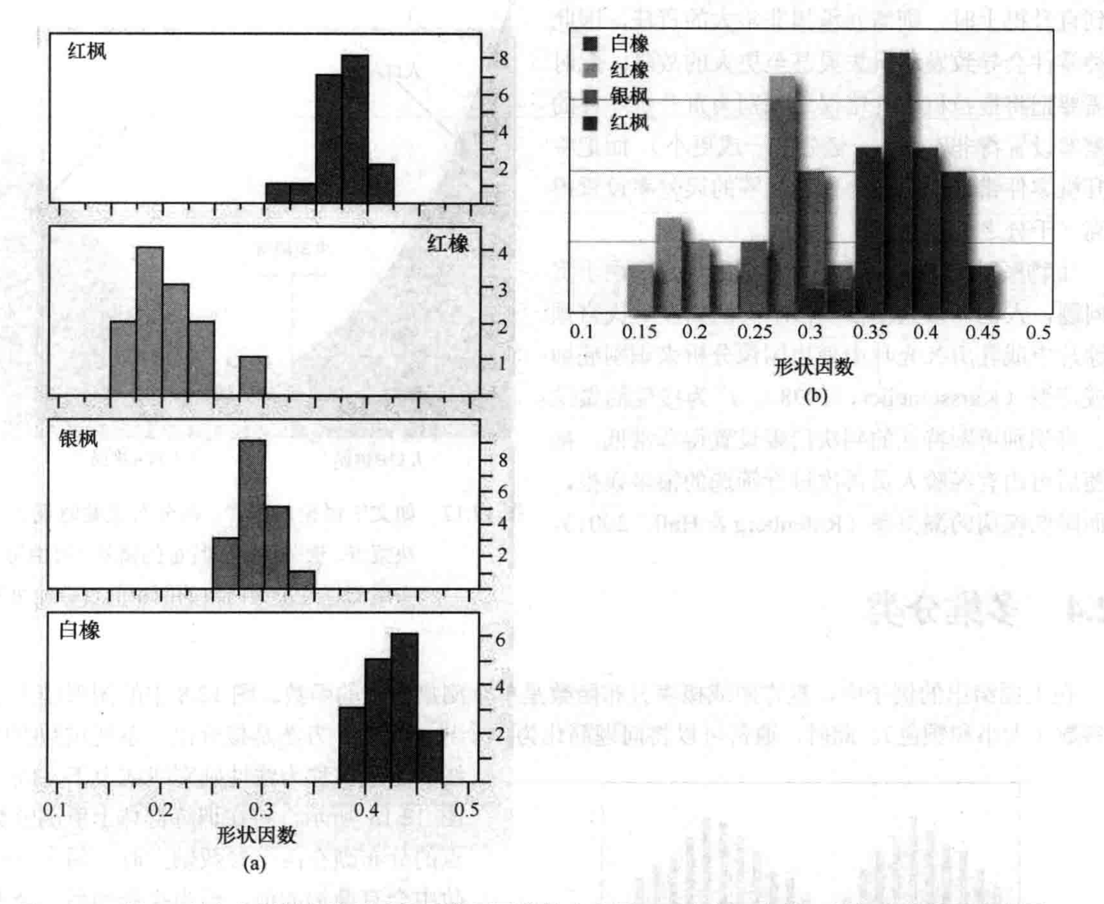


图 12.14 第 11 章中介绍的几种树叶的形状因子值的直方图：(a)红枫、红橡、银枫和白橡的直方图；(b)枫叶和橡叶的组合直方图

引入第二个测量参数——半径比，可解决这一问题。这个参数是最大内切圆直径与最小外接圆直径之比，且与由特征的面积和周长计算出的形状因子无关。这两个参数的线性组合可产生正则最佳坐标轴，在这种坐标轴上，4 种树叶是明显分开的，如图 12.15 所示。

原理上可将非线性回归线用做轴线，但实践中很难进行鲁棒的非线性回归，除非事先已知函数相关的形式，此时事先变换测量变量会更好。例如，不使用面积而使用等效直径（它随面积的平方根变化）来得到更好的线性回归拟合图可能会更好，如图 12.16 所示。

将新轴拟合到数据点后，可使用沿新导出参数投影的值的直方图（如图 12.13 所示）来进行分类，就像使用直接测量的参数来分类一样。例如，图 12.7 中使用的形状因子本身就是面积与周长测量的组合，因此该过程一步就可得到导出的参数来直接区分感兴趣的类。

使用沿两个群体之间的最佳拟合线确定的决策点，可生成一条分开两个群体的决策线（或在高维空间中使用两个以上的参数时，生成一个决策平面或决策表面）。决策线方法可推广到两个以上的类，如图 12.17 所示。此时，两个测量参数（一个是大小参数，即内切圆的半径；另一个是形状参数，即内切圆与外接圆半径之比）足以区分 5 类坚果，如图所示。当应用到（将坚果放在平板扫描仪上得到的）图像时，这一模型成功标记了每个物体。在该例中，参数空间分成了多个矩形。在更普通的情形下，经常使用多边形。下面介绍一种使用到椭圆形区域的距离来定义每个类的不同方法。

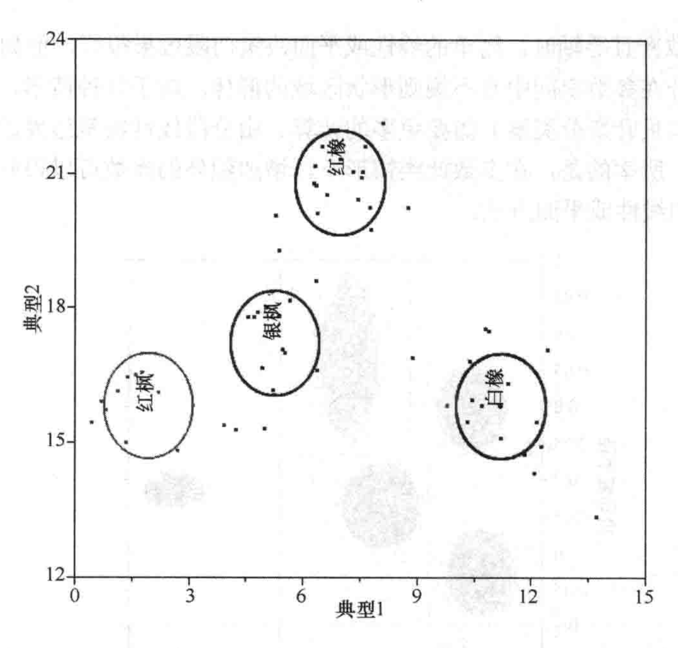


图 12.15　将图 12.14 画在形状因子和圆度的组合轴上，分离了 4 类树叶（线性判别分析用 SAS JMP8 软件进行）

(a)

$R^2 = 0.45$

形状因数

0.0016　面积/μm²　5.10

(b)

$R^2 = 0.51$

形状因数

0.040　等效直径/μm　2.26

(c)

图 12.16　颗粒的扫描电子显微镜图像：(a)形状因子与面积的回归图；(b)形状因子和等效直径的回归图；(c)形状因子和等效直径的回归图更好地拟合了数据

当各个群体非常致密且等轴时，简单的线性或平面决策门限效果很好，但如图 12.18 所示，它需要更高的复杂度来分开在参数空间中有不规则形状区域的群体。对于这种情形，可以找到最佳决策平面或表面的那些技术（贝叶斯分类器）需要更多的计算。由分段线性决策边界或由二次或更高次幂边界也可区分多个群体。所幸的是，在多数此类情形下，增加额外的参数可以很好地分开不同的群体，因此可以使用更简单的线性或平面方法。

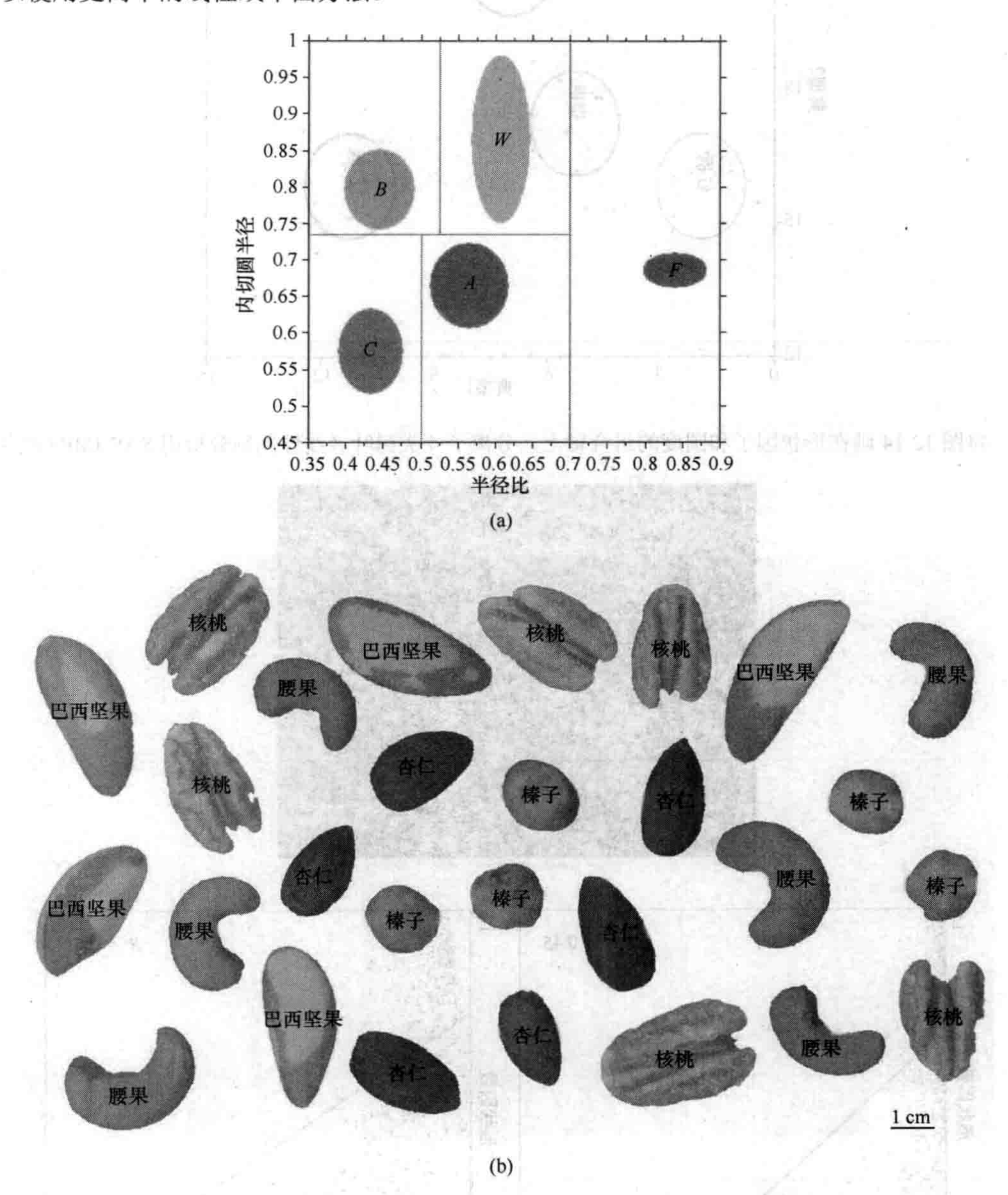


图 12.17　双参数分类：(a)基于内切圆半径与外接圆半径之比，训练群体来识别巴西坚果、核桃、榛子、腰果和杏仁所得到的分类；(b)使用分类方案自动对这些坚果着色和标记。这些类可由参数图上着色的椭圆区域或红色决策线描述

像图 12.18(b)中所示的那些类（它们被拉长并以某个角度面向由测量参数定义的轴）一样，群体类揭示了参数之间的相关性。如上所述，这会引发一个分类问题，因此需要使得群体区域更等轴和更圆。数据集的协方差定义为

$$c(i,j)=\frac{\sum_{k=1}^{n}(x_{k,i}-\mu_i)(x_{k,j}-\mu_j)}{n-1} \tag{12.2}$$

式中，x 是测量值，i 和 j 标识这两个参数，μ 是每个参数的均值，而 k 贯穿群体中的 n 个特征。协方差会在$+\sigma_i\sigma_j$和$-\sigma_i\sigma_j$之间变化，其中 σ 是每个参数的标准差（也可使用归一化协方差，即协方差与两个标准差的乘积之比）。$c(i,j)$为零表示不相关，为最小值或最大值则表示完全相关，如图 12.19 所示。

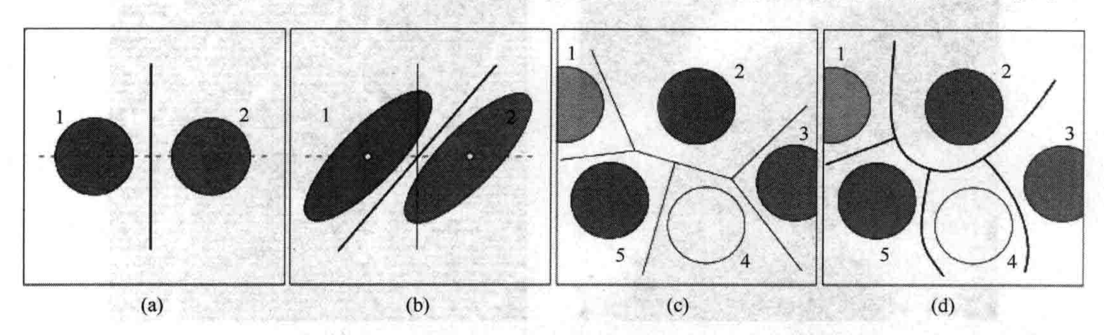


图 12.18　二维情形下的分类示例：(a)分开的各个类表明决策边界与线性判别或回归线正交；(b)未被图(a)中的边界分开的类，使用一个最优贝叶斯分类器可分开它们；(c)被逐段线性决策边界分开的多个类；(d)被二次决策边界分开的多个类

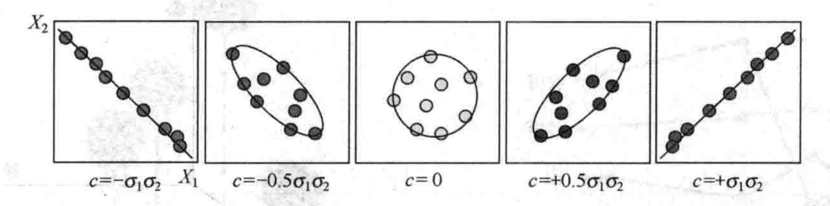


图 12.19　如在文中所讨论的那样，协方差的值会在完全正相关和完全负相关之间变化。零值表示不相关的值

上面的例子表明只有两个测量的或合并的参数，但通常在高维空间会有 n 个轴。此时，所测量的各个参数的所有协方差 $c(i,j)$都可归并到协方差矩阵 $\boldsymbol{C}$ 中。使用这个矩阵可把不规则群体聚类变换成等轴的圆形聚类。对于每个测量矢量 $\boldsymbol{x}$（由 n 个测量参数定义），r 计算为

$$r^2=(\boldsymbol{x}-\mu)'\boldsymbol{C}^{-1}(\boldsymbol{x}-\mu) \tag{12.3}$$

它是从表示特征测量参数的点到群体均值 μ 的马氏距离。这是常用距离概念的一个推广，适合于各个轴有不同含义且相关的测量。在等轴群体聚类的情形下（协方差为零），马氏距离即欧氏距离或毕达哥拉斯距离。

通过将距离测量归一化，沿各个轴测量距离有助于补偿不同比例尺的各个测量参数。这是通过将每个距离除以一个适当的比例常数来实现的，比例常数通常是参数值的标准差。这称为标准化的欧氏距离。

当然，分类方法可推广到二维以上的情形（此时使用简单的图形来显示会更困难）。图 12.20 显示了遥感成像的一个例子，其中使用了在地球资源卫星专题测绘仪的不同频带中检测的反射光强度来对地形分类（Grasselli，1969；Sabins，1986）。频谱强度图案与特定类型的植被、矿物等相关。使用了一个监督训练程序，其中一名操作员标记位置，而图像分析系统则根据在每个波段中测量的强度来画出那些点。然后聚类点标记不同的类，进而使用不同的类来逐点标注图像。

当然，我们没有理由期望与每个类对应的聚类点在 N 维参数空间中形成一个规则的球形，下面的几个例子给出了建立这些区域的其他方法。最简单的方法之一是为每个参数值设立上限和下限。这相当于参数空间中定义每个类的一组矩形框（见图 12.21）。因为无须找到一个有代表性的群体，而只需

找到或预测极值，因此训练这样一个系统很简单。

建立区域边界的另一种方法是，对每个类测量一个有代表性的训练群体，并由每个分类参数的均值和标准差来表征值的分布，如上面的图 12.15 所示。使用均值作为中心坐标，并使用每个参数的标准差的倍数作为轴的长度，就会形成椭球体，如图 12.22 所示。这些椭球体的表面可用做绝对的类边界，但它对按标准差来测量每个新特征参数坐标到各个类的距离更有用。

图 12.20　使用地球资源卫星专题测绘仪图像进行土地利用分类：(a)显示了单个波长范围内的反射率的图像；(b)反射率与波长关系图，以及不同地形反射率值的聚类图；(c)基于这些数据的土地利用分类

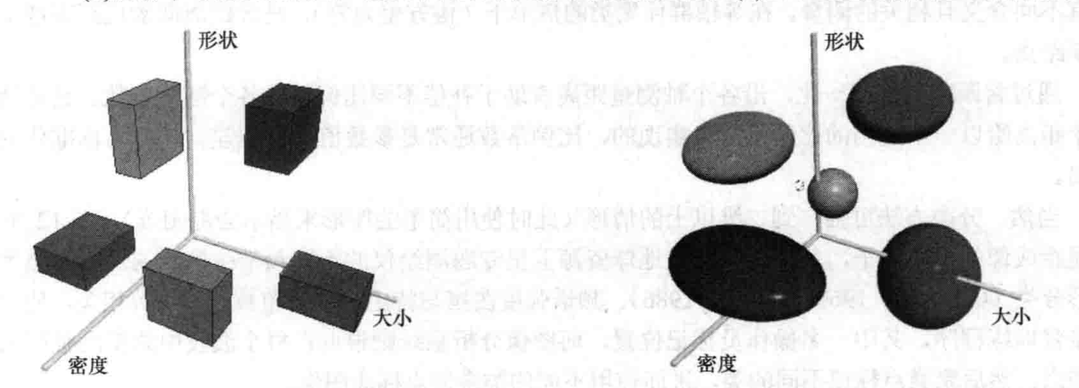


图 12.21　参数空间中类区域的限制范围图（说明了 3 个独立的参数）。这是图 12.17(a)中红线框所示例子的一种三维形式

图 12.22　由测量数据的标准差计算的椭球体类区域。这是图 12.17(a)中彩色椭圆形区域所示例子的一种三维形式

为简单起见，图 12.23 显示了一个二维示例。各个训练群体显示在顶部，下方显示了点图及每个类的 2σ 椭圆图。注意，对于训练群体和随后测量的其他特征，有些点不出所料地位于范围 2σ 外部。另外，有些类要比其他类大得多（参数值的更大变化），且由于一个参数比另一个参数变化更多，椭圆可能会被拉长。由每个椭圆的轴的长度显示的标准差，会归一化各个轴的单位，因此成为了一个用于解释参数空间中其他混淆距离的比例。参数空间中的“距离”概念并不清晰，它在不同方向或图形的不同区域可能会失真。遗憾的是，使用形状因子进行识别时，情形通常如此，因为它们通常不都是独立的（因此各个轴不正交），进而也必定不是线性的（例如，长宽比从 1.1 变化到 1.5 远远不同于从 4.1 变化到 4.5）。

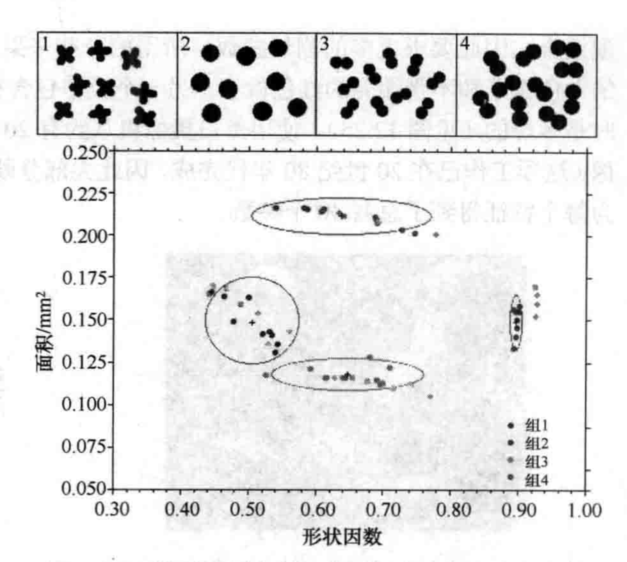


图 12.23 统计范围内的二维示例。注意，有些点在 2σ 边界之外。实心点表示显示在顶部的训练群体；灰色点表示识别的其他特征。如文中讨论的那样，点以最近的聚类识别

如果测量了一个新特征而其参数坐标未放置在一个类边界的内部，那么通过在参数空间中测量特征坐标到每类中心的距离（均值），并通过将这些距离除以标准差来进行标定，它仍可识别为“最可能”的一类（平均值）。由于聚类的大小不同，这就意味着一个点可能更靠近图上的某个类，但却更“像”一个不同的类。在该例中，靠近紧凑蓝色类的一个点与较大的红点相比，标准差可能更小，因而可据此进行识别。

12.5 学习系统

收集真正有代表性的一组特征来建立训练群体，进而使用这些训练群体来产生各个类而不改变许多新特征和情形，是非常困难的。特别地，靠近分布两端的极值特征通常很少。此外，使用均值和标准差而非实际形状分布的方法可能会使测量过于简化，在许多图像分析情形下，这种方法并不产生只由这两个统计参数就能充分描述的高斯分布。

上面的图 12.12 说明了由形状分布直方图来估计分类误差概率的能力。因此，使用实际的直方图可扩展传统的“硬”分类，此时可将一个特征分配给某个类，或从该类中排除这个特征，还可将一个特征置于“模糊”逻辑情形。模糊逻辑是经典逻辑的重要延伸，其知识库中的各种分类规则包含概率。例如，“若 X 会飞，则其为鸟的概率是 70%；若 X 不会飞，则其为鸟的概率只有 5%”，这包含了蝴蝶（会飞但不是鸟）和企鹅（是鸟但不会飞）。在参数测量情形下，值的直方图可直接测量条件概率。

模糊逻辑已应用于专家系统和神经网络（两者都在下面讨论，它们是基于多个规则来分配类的工具）。通常这会减缓收敛过程，因为有必要评估所有路径并找到最有可能的路径，而不只是找到一条满足这些规则的成功路径。在正式意义上，模糊逻辑等效于具有多层次经典集合的贝叶斯逻辑。关于模糊逻辑的综合信息，请参阅 Zadeh（1965），Negoita & Ralescu（1975，1987），Sanchez & Zadeh（1987），Zimmerman（1987）和 Bishop（1996，2007）。

使用直方图来建立决策点，并采用线性判别线或上下文线方法在多参数空间中运作的学习系统的一个示例，使用了 9 类坚果的群体（见图 12.24）。与图 12.17 中的例子相比，这是一个更困难的坚果类

型组合，因此要求更多的测量参数。所用的一些坚果代表了多个类别（例如，一个类中包含有带端壳的白色橡子和不带端壳的红色橡子，另一个类中包含有几种胡桃），因此直方图一般不是高斯的，且有时是多峰的（见图 12.25）。使用黑白摄像机从约有 20 个不同坚果的一个初始训练群体中获取了几幅图像（这项工作已在 20 世纪 80 年代完成，因此大部分硬件按当前标准是简单粗糙的），并且测量了特征，为每个特征得到了总共 40 个参数。

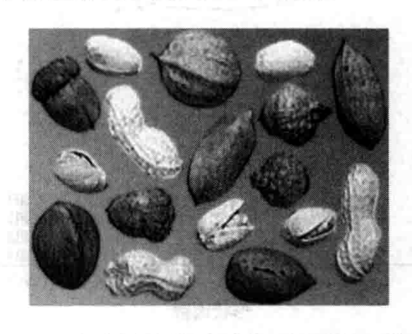


图 12.24　文中讨论的用于学习例子的一些坚果

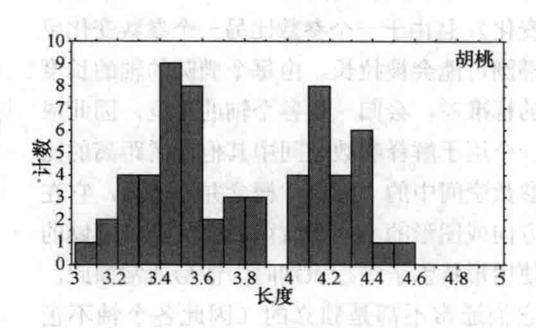
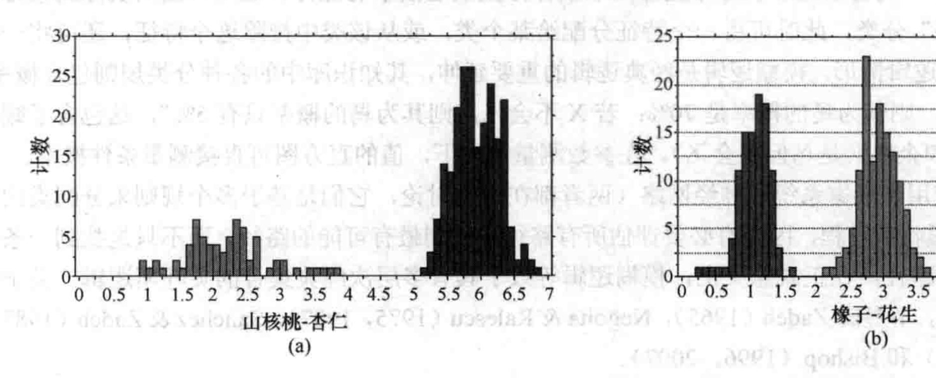


图 12.25　两种胡桃的双峰长度分布

有些参数，如位置和方向，可由人类判断它们没有意义而被丢弃。随后的统计分析丢弃了其他一些参数，只留下了 17 个实际使用的参数，因此辨识在 17 维空间中进行。但并非所有维度都真正是独立（正交）的，因为许多形状参数基于各种组合中使用的尺寸测量。因为在该空间中，线性判别线是在每对坚果类别之间采用逐步回归构造的（Draper & Smith，1981）。基于 F 值（使用了一个任意截止值 4.0）将参数添加到上下文线的方程中，或从方程中去掉上下文线，消去了那些不改善区分特定群组性能的变量。

这种变量的消去有助于达成使用独立参数的目标，且有助于简化整个问题并减少所需的训练示例数量。例如，许多形状参数使用一些相同的尺寸信息。譬如第 11 章中定义的凸度和形状因子二者都使用周长。如果其中一个参数与类别区分高度相关，那么包含周长的其他参数也会有很大的机会显示一种相关。逐步回归可以选取具有最高相关和预测能力的一种参数而丢弃其他参数。

生成的上下文线最少使用两项，最高使用八项。有些参数可用在多种情形中（如平均亮度），而其他参数只出现在一个或两个方程中，但它们通常是最重要的单个参数（如凸面积在区分开心果和杏仁时是非常重要的）。平均而言，在总共 36 个成对的上下文线中，这 17 个参数中的每一个会使用 11 次，而每个上下文线方程平均使用 5 个参数。图 12.26 显示了几个例子。线性判别线或上下文线方程中的系数是通过回归导出的，并且表示了该线与 n 维空间中每个参数轴之间的角度。



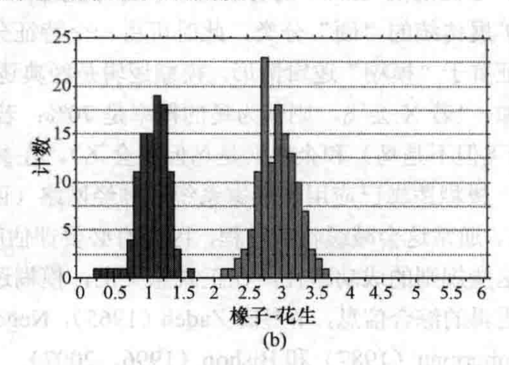


图 12.26　山核桃－杏仁和橡子－花生的上下文线与频率分布，完全分开各个类的例子。回归方程如下：山核桃（类型 2）－杏仁（类型 6）= −2.266 − 2.372 幅度 + 2.184 宽度 + 0.226 亮度 + 0.318 对比度 + 2.374 纹理；橡子（类型 1）－花生（类型 3）= 3.928 + 0.773 长宽比 − 5.007 分形维数 + 0.068 亮度

从杏仁中区分山核桃以及从花生中区分橡子很容易，因为它们的分布完全不同，尽管在形状上这些分布有点不规则。橡子－胡桃的分布（见图 12.27）稍有重叠，它约占成对判别图中的 16%。使用图 12.12 中所示的逻辑，决策点位于误分类误差概率对两个类相等的位置。在橡子－胡桃情形下，该数量约为 1%，且在从侧面观察一个没有端壳的大橡子时出现，如图 12.28 所示。最可能的错误是混淆橡子与榛子。

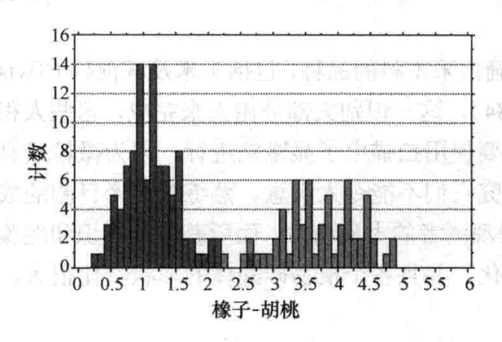


图 12.27　橡子－胡桃的上下文线和频率分布，显示有重叠。回归方程：橡子（类型 1）－胡桃（类型 4）= 1.457 + 6.141 形状因数 – 2.153 凸周长 +1.768 凸区域 + 2.926 长度 – 1.575 幅度 – 1.131 范围

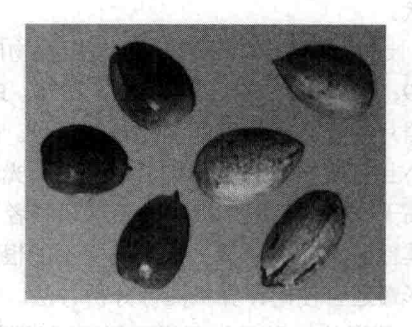


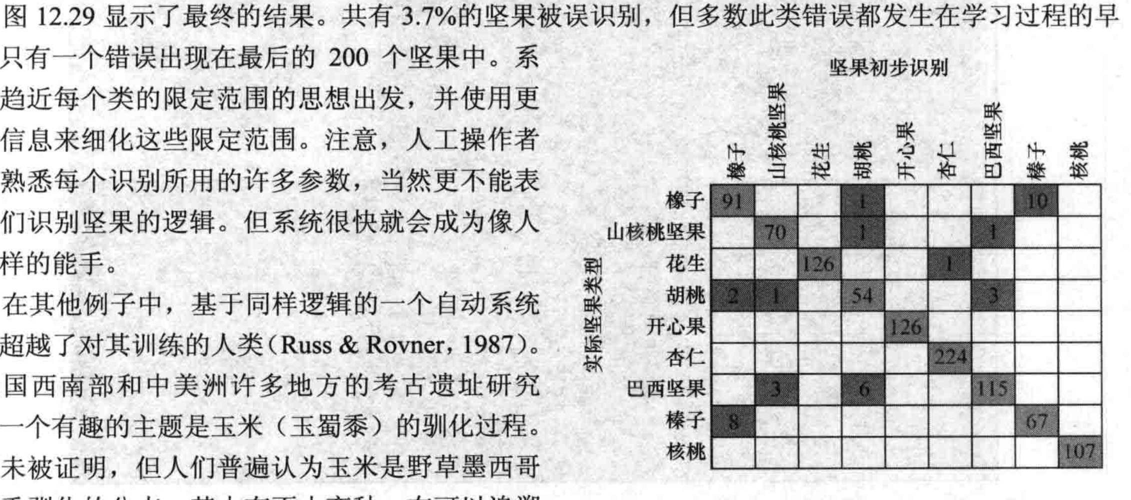
图 12.28　观察几个没有端壳的大橡子和胡桃，显示了两类之间可能出现混淆的一种情况

建立判别函数（上下文线方程）并沿这些线建立参数值的分布后，系统就可识别坚果。沿每条上下文线出现的决策点指定几个平面（与这些线正交），在高维空间中这些平面围绕每个类形成多面体。问题是，使用一个来自每个类的仅有 20 个坚果的训练群体，可合理地为每种坚果类型建立聚类的位置，进而充分定义上下文线，但决策点的位置并不完美。

由于系统会检查更多图像（最终的总数会超过 1000 个坚果），因此会计算导出参数值（沿每条上下文线的位置）并更新直方图。由于已建立了直方图，因此这并不要求额外的计算机内存。根据存储的决策点，继续识别坚果类型。但在将一个值添加到直方图末端（其面积的最后 5%以内）时，程序会要求操作员确认这一识别，然后使用“过去－现在”直方图重建决策门限，使得每类错误的概率相等。采用这种方式，系统由已知的额外标本开始学习，在每个分布的重要界定区域中逐渐增加观察的数量。

图 12.29 显示了最终的结果。共有 3.7%的坚果被误识别，但多数此类错误都发生在学习过程的早期，只有一个错误出现在最后的 200 个坚果中。系统从趋近每个类的限定范围的思想出发，并使用更多的信息来细化这些限定范围。注意，人工操作者并不熟悉每个识别所用的许多参数，当然更不能表示他们识别坚果的逻辑。但系统很快就会成为像人类一样的能手。

在其他例子中，基于同样逻辑的一个自动系统已经超越了对其训练的人类(Russ & Rovner, 1987)。在美国西南部和中美洲许多地方的考古遗址研究中，一个有趣的主题是玉米（玉蜀黍）的驯化过程。尽管未被证明，但人们普遍认为玉米是野草墨西哥类蜀黍驯化的分支，其中有不少变种。在可以追溯的考古学文献中几乎没有关于玉米的信息，但巧合



坚果初步识别（列）／实际坚果类型（行）

实际坚果类型	橡子	山核桃坚果	花生	胡桃	开心果	杏仁	巴西坚果	榛子	核桃
橡子	91			1				10	
山核桃坚果		70		1			1		
花生			126			1			
胡桃	2	1		54			3		
开心果					126				
杏仁						224			
巴西坚果		3		6			115		
榛子	8							67	
核桃									107

图 12.29　显示初步坚果识别成功（绿色）和错误（红色）的“混淆矩阵”

的是，像所有禾本类（和许多其他植物）那样，玉米的茎、叶和其他部分的细胞内和细胞间出现了乳白色的植物硅酸体。这些硅酸体的大小通常只有几微米，作用是使植物组织变硬。它们还会选择性地出现在植物的损伤部位。

从人类学的观点来看，我们对植物硅酸体感兴趣的原因有两点。首先，它们能在土壤中存活很长时间，能从可确定其年代的土层中恢复，并表现出人类居住的其他迹象。其次，它们具有与众不同的形状。

过去几十年的研究表明，使用植物硅酸体可以确定禾本科的品种，包括玉米及其前体（Twiss et al.，1969；Rovner，1971；Pearsall，1978；Piperno，1984）。这一识别大部分由人来完成，这些人积累了许多照片并建立了可供匹配比较的目录。这项工作主要使用扫描电子显微镜进行，因为微粒只有几微米大小且有明显的三维结构，所以使用光学显微镜研究它们不能令人满意。熟练观测者目测的成功率一般高于 95%（仅限于很少的熟练观测者）。这会耗费观测者的大量精力，而所获取的知识和经验也不易于其他研究人员掌握。人眼确实不能很好地描述变化，并且各个植物硅酸体的形状变化很大。这样的情形最适合使用计算机图像分析方法。

使用从 5 种已知玉米和墨西哥类蜀黍的一个训练组中提取的植物硅酸体，制备了几个载玻片。对 SEM 图像进行分析确定了几个大小和形状参数，并按上述方法使用这些参数建立了判别类。对原始的目标集，达到的识别率高于 99%。对来自另外 5 种植物（3 种玉米和 2 种墨西哥类蜀黍）总共 300 个目标应用这一结果后，识别率超过了 99.6%。此外，该系统 1 个月就可以从各类玉米或基他草本植物中检出几百种植物硅酸体（足以产生统计学意义的结果），相当于过去 20 年通过手选和目测检出的总和。近来，该系统一直用于南瓜种子的分类（Rovner，个人通讯）。

这些例子均基于对象的群体，并建立了涵盖生物标本中自然出现的变化的各个类。一个更加雄心勃勃的项目试图通过识别和测量个体足迹的关键尺寸与形状参数，来识别类中的个体。图 12.30 显示了几种野生动物的足迹。足迹识别技术（FIT）是监测濒危野生动物物种的一种成本低廉且不对动物造成影响的方法。

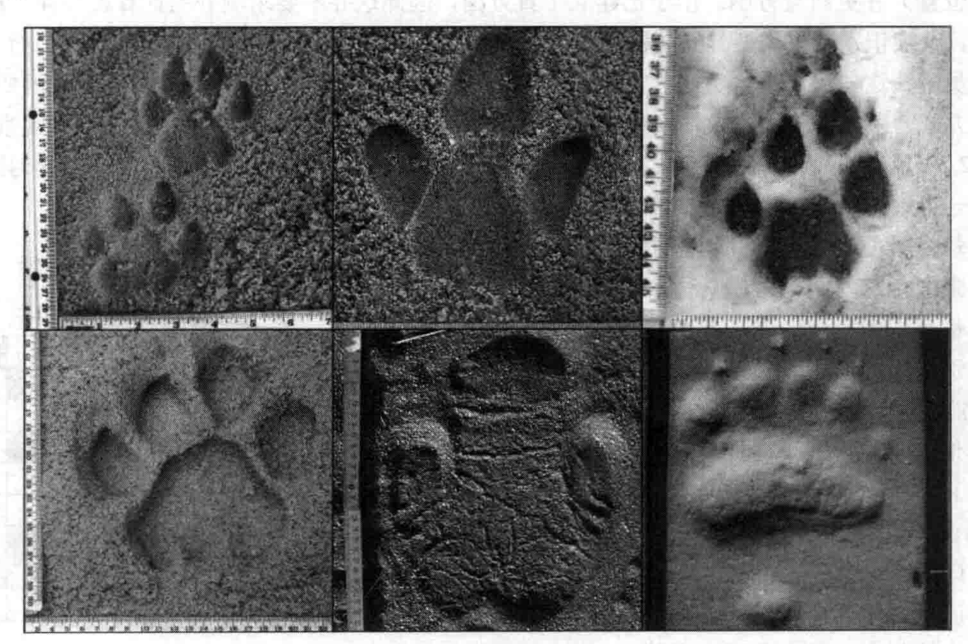


图 12.30　用于识别野生种群中个体动物的足迹（从左上角起：美洲狮，貘，猎豹，虎，白犀牛，北极熊）。如标尺所示，图像的比例尺是不同的。踪迹在沙子、泥土和积雪中记录

全球生物多样性正面临着前所未有的危机，目前世界上多达25%的哺乳类动物濒临灭绝的危险。收集这些物种的数量和分布的数据是制定有效保护政策的基础。跟踪，即利用足迹来识别并追寻动物，可能是世界上历史最悠久的行业之一。FIT是一种客观的科学方法，它可把古老的动物跟踪技术应用到野生动物的监测之中。侵入性方法之后采用的麻醉并安装无线电项圈的方法，已被证明降低了黑犀牛的生育率，而在识别动物足迹形状方面，FIT方法已表现出了很大的灵活性。已对黑犀牛和白犀牛、南美貘和中美貘、北极熊、美洲狮等改写了算法，对其他物种的算法正在开发之中（Alibhai & Jewell，2001，2006，2007，2008）。

每种动物都有其独特的足迹，各物种内的每个个体也都有不同的足部结构（类似于人类的指纹或面部结构）。此外，每次脚在基质（可能是灰尘、泥土、墙面板、积雪等）上留下的足迹都不相同。这三种层次的变化（物种、个体与基质）就导致了足迹识别的复杂场景。动物的步态和质量、风力和风向、光照等影响足迹的因素，会使问题更为复杂。为试图包含一些此类变化，人们通过收集这些物种中的15～20个“已知”个体（通常是圈养动物）的15～20个足迹，生成了初始的物种FIT算法。

足迹的数字图像都是在野外拍摄的，对图像处理后，通过测量长度、角度和面积，画出了足部的几何剖面。目前，测量是手工进行的，但有望在未来自动化。使用JMP软件（SAS Institue，Cary，NC）选取了以要求的级别来进行辨别的变量。多元统计分析和建模过程都可按精度等级来对圈养动物和野生动物分类，个体的识别率通常在95%以上，且FIT还可对物种、年龄段和性别进行分类。FIT程序可以提供普查数据来确定一个未知群体中的动物数量，并对个体进行跟踪监控。由这些数据可给出群体估计和密度分布，进而为制定野生生物保护政策提供数据支持（WildTrack.org）。在许多方面，这类似于识别监控图像中的个体。打印图像中少量的参考点和基质引入的变化，会使得这种识别更加困难，但因匹配的数量较小而更简单

12.6 *k*NN和聚类分析

在参数空间中确定类限还有其他方法。最广泛使用的两种方法是k近邻法（kNN）和聚类分析，但与上面描述的直方图和线性判别方法不同的是，这两种方法都有一个缺点，即必须为每个识别的特征保存实际的n维坐标（对多个参数的测量值），这意味着如果系统继续学习，存储需求就会不断增加。这两种方法的另一个缺点是，随着所存储的值的数量的增加，识别所需的时间也会增加，甚至使用高效算法排序或裁剪数据集来将搜索限定到合理数量的候选点上时，也是如此。另一方面，这些方法并不假定类别区域具有任何特殊形状，如上面描述的长方体、椭球体或多面体方法。甚至可能会处理如图12.31所示的情形，即一个类可能是非凸的、不相交的，或被另一个类部分包围或完全包围。

k近邻分类的工作方式是，搜索先前所识别特征的数据库，找到与当前测量特征“最像”的那些特征，如图12.32所示。这相当于求当前特征参数的坐标与其他特征参数的坐标之间的距离。该距离通常在毕达哥拉斯意义下计算（即差值平方和的平方根），但这忽略了一个重要事实，即参数空间的不同轴具有不同的单位和测度。沿每个轴的距离通常表示为该参数数值总范围的一小部分，但这样假设的理由并不充分。最好使用一种归一化尺度，如每个类内各个值的标准差（对马氏距离可使用上面椭球体类限的相同处理方法）。

在最简形式的k近邻比较中，k等于1，且搜索只针对最相似的单个现有测量，接下来假定用它来识别新特征的类。对于靠近类间边界的特征，单个最近邻匹配已被证明是相当带噪的，并会产生不提升识别鲁棒性的不规则边界，如图12.33所示。使用大多数5近邻、9近邻甚至更多近邻进行识别，会产生更平滑的边界，但仍对群体的相对大小高度敏感。如果某个群体中的成员比另一个群体中的成员

多得多，则会有效地压缩较小群体的类别范围，方法是使得大多数匹配更可能是与更大群体的成员匹配。此时就强烈地排斥了仅代表总观测数量几个百分点的小群体。

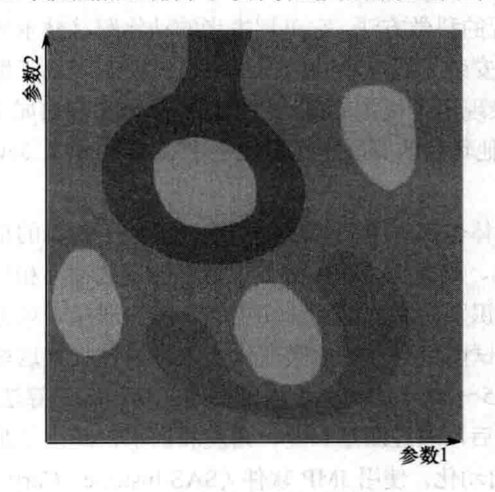


图 12.31 简单的几何形状难以表示的非凸、不相交或包围类的例子（用不同颜色标出）

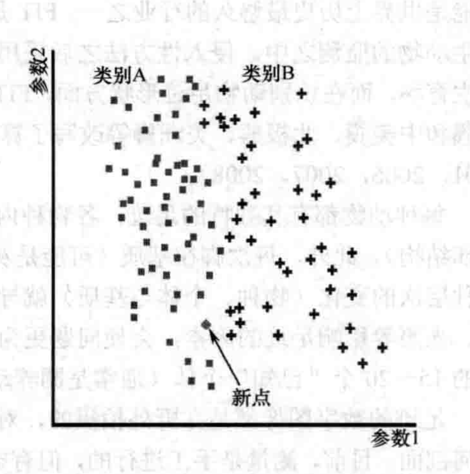


图 12.32 k 近邻分类：新点的 5 近邻中有 3 个位于类 B 内，因此该点分给类 B

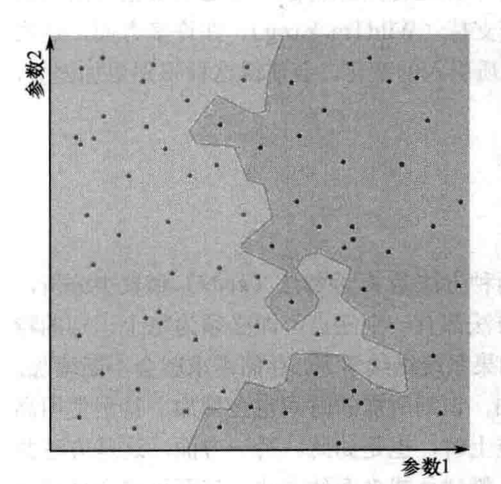


图 12.33 使用单个近邻分类时，在两个类（各个成员的参数值由不颜色的点表示）之间形成的不规则边界

迄今所描述的方法均假设可使用一些训练群体（其中的特征属于已知类）来构造类别范围。这就是“监督学习”，学习过程中，由操作人员识别每个特征应归到的实际类或群体（基于任何其他信息、独立的人为判断或可用的先验知识），而软件使用这种信息以及可用于确定一种识别策略（这种策略通常完全不同于人类所用的策略）的任何测量参数。有些情形下，事先并不知道这些类是什么，甚至不知道会出现多少类。此时，在参数空间中画出每个特征的测量参数值，然后查找聚类点，仍是有用的。非监督训练聚类分析这一主题非常丰富，完全超出了本书的范围（Pentland，1986；James，1988；Bow，1992；Fukunaga，1990）。

聚类分析处理的一种典型方式是，使用点间的距离来识别聚类中的成员。例如，从靠得最近的两个点开始。假设这些点表示一个类，并将其他近邻点加到这个类中，直到已分配给该聚类的最近点的距离超过某个（任意）限度，这个限度可由不断增长的类区域的尺寸或统计计算出来（如超出 2 个或 3 个标准差）。该方法可推广到使用一个以上的近邻。给定类停止增长后，再次从剩下的点开始并继续上述过程，直到这些点都归到不同的聚类中。这种方法有许多改进版本，它们可合并或分裂所分配的各个类区域。

寻找聚类的一种互补方法是同时使用所有的点。它从为所有出现的点构建一颗最小生成树开始（对于大型数据集，这需要大量的计算）。这棵树上的链可由其欧氏或马氏长度加权，或由每个点的 k 近邻表中出现的匹配数加权。换言之，如果这些点都列在彼此的近邻列表中，那么它们可能在同一聚类中。然后使用最大长度或最小的匹配数来修剪这些链，进而修剪生成树，以留下识别出的各个聚类。图 12.34 显示了点的一个稀疏聚类、最小生成树和修剪后得到的聚类的一个例子（Bow，1992）。修剪过程将根

据一个指定数量的聚类或剩余链的最小长度来中止。

在数据中查找各个聚类的另一种稍微不同的方法是树状图，尤其是在只有一个重要的测量变量或几个（由逐步回归或线性判别分析导出的）变量的组合的情形下。图 12.35 显示了一个例子。特征已按测量变量的数值排序，比如等效直径或一些正则变量（观测变量的组合）。然后用水平线连接那些具有最小差值的特征，而水平线的垂直位置对应于值中差的大小（本例中的数字 7 和 21，接着是 4 和 23、27 和 14、13 和 16 等）。特征组通常基于组中任何两个成员间的最小差来链接，但也可使用每个组中的均值或中值之差。最终，所有特征和组被连接为一棵树。

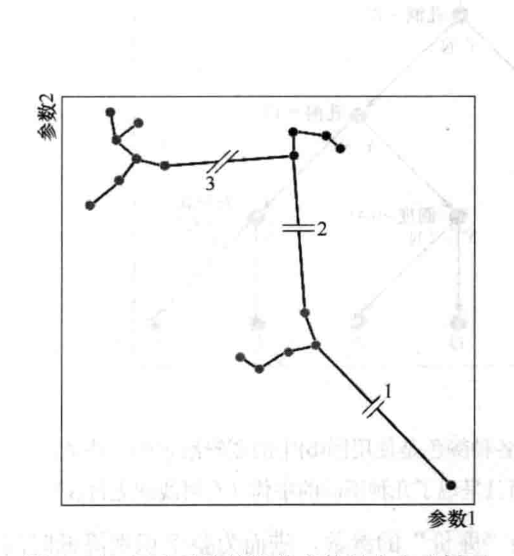


图 12.34　以点及连接这些点的最小生成树画出的参数值。使用最大长度切割分支分开了聚类（用不同颜色表示）

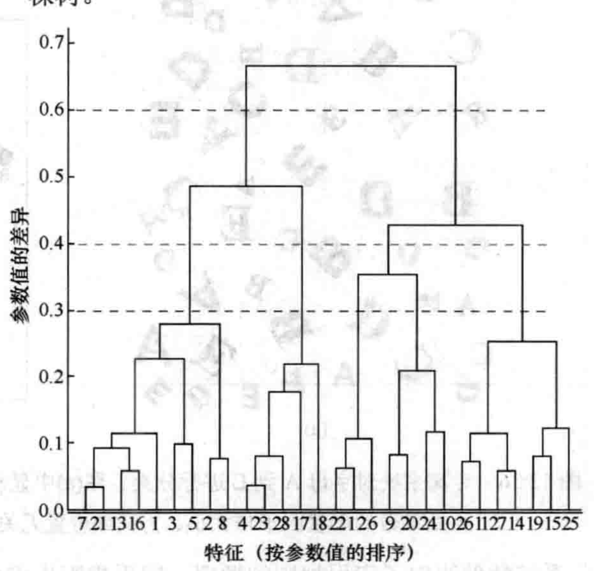


图 12.35　文中讨论的树状图示例。红色、绿色和蓝色线表明了为特征测量间的差选取不同阈值来建立分类数的效果

这种结合了实际测量参数和它们的排列顺序的方法，是一种典型的非参数分类和统计比较方法。这意味着它们不依赖于统计特性（如假定值为高斯分布的均值和标准差）。如果使用多个独立的测量参数（如形状、大小和颜色），则特征之间的差计算为 n 维空间中表示多个值的矢量间的距离。这通常会引发一个问题，因为各个参数的单位完全不同，且各个轴的长度也不协调，故没有一种自然的方法来按比例缩放它们。有时，基于测量参数的总范围或标准差来归一化每个测量参数，可解决该问题，但这种方法随意性太强。若不能完全表征每个参数的统计分布，就不能为参数空间计算真实的马氏距离。

构建树状图后，就可用它来求任何聚类数，方法是为差选取一个不同的阈值。例如，在图 12.35 中，阈值 0.6 产生 2 个聚类，阈值 0.4 产生 4 个聚类，阈值 0.3 产生 5 个聚类。由于在无监督聚类情形下按定义事先并不知道会出现多少个类，因此门限值（或聚类数）的决策需要人工判断。

12.7　专家系统

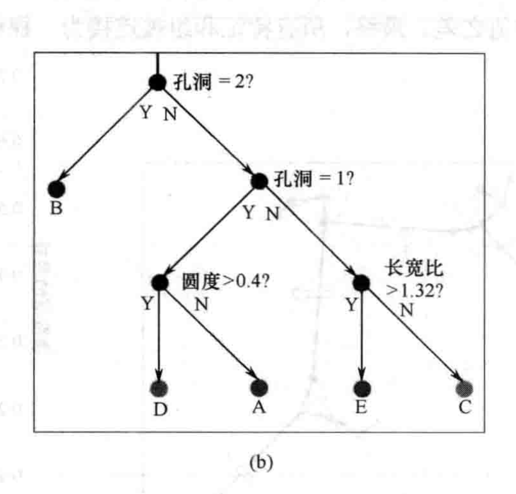
在通过某种方法识别聚类后，仍需要使用上面所示的方法之一来构建类限。它可以是一种几何形状（如长方体、椭球体或多面体）或一条 *k*NN 边界。接着需要考虑如何应用这些类边界作为一组规则来识别后续的特征。

最简单和最常见的方法是传统的专家系统。创建一个规则集（如类边界的限制——要么硬要么模糊），然后将它应用到每个新的特征测量组。回到字母 A 到 E 的例子，图 12.36 显示了一个非常简单的专家系

统，它只有 4 个规则，且只需要测量 3 个参数。在最有效的实现中，只有那些需要的参数才被测量，所以对于字母 B（它由形状中的孔洞数识别），完全不能确定根本不需要的圆度和长宽比参数。通过测量由人类“专家”指定的几种不同字体的字符，对每种字型，用实验方法求出了规则中参数的决策值。



(a)



(b)

图 12.36　专家系统对字母 A 到 E 进行分类。图(a)中显示的各种颜色是使用图(b)中的逻辑给定的。注意，用做参数的形状因子与大小、方向或位置无关，而且呈现了几种不同的字体（有衬线或无衬线）

系统软件决定了应用规则的顺序。应用规则先求出最“廉价”的参数，进而为整个识别降低时间和精力通常非常重要。这就要求每个规则都有与其相关的成本。统计孔洞数量与测量形状因子的参数相比，计算量会少一些。还要注意的是，与上面显示的模板匹配和互相关例子不同，这种方法对字母的位置、大小或方向不敏感，并能容许几种不同的字体。

第 11 章中的图 11.35 表明识别字母 A 到 E 这个特殊的问题，可由简单无衬线字体的端点、节点和环的拓扑性质来完全解决。通过可控的修剪，删除此处所用字体中的衬线（即删除短的末端分支），允许使用相同的过程；所示方法说明完成给定任务可能有几种方式。

这样的简单分类系统有时也被称为决策树或产生规则，它由一组规则的 IF ... THEN 关系组成。经典专家系统分开了数据库（规则和它们的成本）和用来求解的推理机（Black，1986；Winstanley，1987；Rolston，1988；Slatter，1987；Walters，1988）。这种系统的一个缺点是，加入另一个类（如字母 F）并不只是嫁接了另一个步骤，而是可能会完全重排应用规则的顺序，甚至会从整个决策过程中删除某些规则并用其他规则替代。最佳决策树不必是具有最少分支的决策树。“知识成形”过程（如包含得到每个输入值的代价）可极大地提升整体效率，并确定一个最佳搜索和求解树（Cockett，1987）。

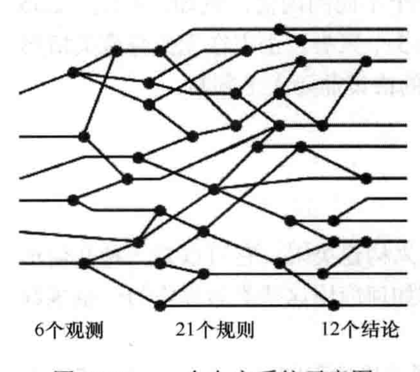


图 12.37　一个专家系统示意图

与这个专家系统相比，多数真实的专家系统的规则要多得多，且应用这些规则的顺序无关紧要。决策路径的复杂性会随着大量规则迅速增加。通用专家系统试图以双向方式找到各个输入值与结论间的一条路径。如图 12.37 所示，如果观测数小于结论数，则使用正向链，它从左侧开始并使用各个规则来探索指向结论的

路径。在其他情形下，使用反向链可能更有效，即从结论向观测反向探索。在这幅示意性草图中，有6个最初的观测（测量）、21个规则和12种可能的结论；真实系统要复杂几个数量级。图中所示的一些规则有两个以上的可能结果；在某些系统中，每种可能的结果用一个单独的规则表示。

对穿过该网络的一条成功路径，正常的前向搜索从左侧的某些点开始，跟随一条路径到达一个规则，并跟随结果，直到出现矛盾（规则未被满足）。系统然后回溯至前一个节点，并尝试一条不同的路径，直到达到一个结论。这种方法并不测试从观测到结论的所有可能路径。对控制可能路径的顺序的试探法进行测试是非常重要的。在复杂的现实中，不可能测试应用一个规则产生的所有后果，因此采用了诸如深度优先或宽度优先的简化策略。还用到了修剪，或在穷尽一条路径前拒绝该路径（作为搜索的类比，我们考虑在国际象棋程序中下一步棋：最初几步可能就会被拒绝，而其他几步则会搜索到各种深度来评估各种可能的反应与结果）。反向搜索的工作方式与此类似，即从可能的结论开始，并搜索通向观测的路径。有些搜寻“引擎”组合了正向和反向搜索方法。

使用模糊逻辑时，各种规则都包含概率。此时，有必要合并每条路径上的节点值来构建一个总概率，以便选取最可能的路径，进而得到最可能的结果。控制搜索顺序的试探法并不是很重要，但总的计算量会非常大。模糊逻辑方法的优势是，可处理不完整的信息，并为每个结论给出一个置信水平或概率水平。

12.8 神经网络

使用专家系统（尤其是使用模糊逻辑概率的专家系统）搜索多条路径，本质上是并行问题。求解该问题的一种较好方法是使用神经网络，即一个简化且理想的生物过程运作模型。神经网络中的每个元素是一个阈值逻辑单元，如图12.38所示。它类似于功能神经元的简化模型，如经典McCulloch-Pitts模型中所示的那样（McCulloch & Pitts，1943；Rosenblatt，1958；Lettvin et al.，1959；Minsky & Papert，1969；Grossberg，1988）。多个输入（在图像分析情形下，它可能会是观测参数值、像素亮度或其他逻辑单元的输出）被加权并求和，并将总和与阈值进行比较。如果超过阈值，就激发神经元。在生物系统中，它发送一串脉冲作为其他神经元的输入；在神经网络中，它将一个计算的输出值（一个概率）传送到另一阈值逻辑单元上的一个输入。

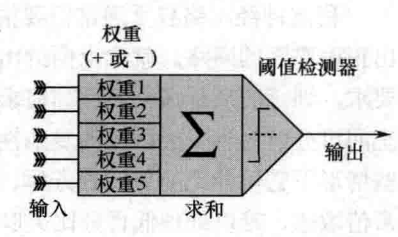


图12.38 一个阈值逻辑单元的图解

为了理解这些简单单元的运作，再次考虑称为“祖母细胞”的理论结构的经典描述是值得的。假设有一个细胞（或一个逻辑单元）已被训练来识别你的祖母，并且会耐心地检查从你的眼睛传来的每幅图像，以搜寻线索。早期的逻辑单元（有些在眼睛中，有些则在视觉皮层中）只查找低级结构。因此，白头发、蓝眼睛、眼睛间距、鼻子、嘴巴和颧骨等，都是祖母细胞的输入。业已证明，人类视觉能非常有效地定位面部特征（甚至刚出生的婴儿也可做到这一点），而且如上所述，这些特征之间的间距之比是识别各个个体的重要线索。

如第2章所述，其他线索可能是熟悉的服饰、首饰或眼镜。有些线索会比其他线索更重要（有更大的权重），而有些线索可能会缺失（如看不到祖母眼睛的颜色）。还有一些线索具有较大的负权重，如胡须浓密或身高超过6英尺。当所有的输入相加后，如果总和足够大（超过一个阈值），那么细胞会通知你的大脑祖母出现了。这样的一个系统会对误报负责。结果并不总是正确的（在某些情形下你可能会错过祖母，或者偶尔会误认他人为祖母），但它非常有效，且运行迅速。

将多个层中的这些简单设备连接成一个网络，会生成一个可基于不同类型输入做出决策的系统。

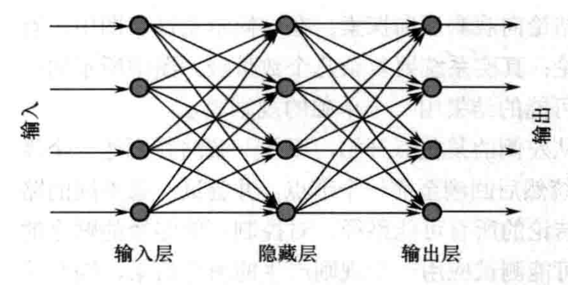


图 12.39 具有一个隐藏层的最小神经网络示意图

神经网络的经典拓扑如图 12.39 所示。每个逻辑单元计算一个输出，该输出是其各个输入值的加权组合（不一定是线性的）。这些输出在到达输出层之前，至少要经过一个隐藏层。输出层产生离散的输出，选择一个或最多几个值。对神经网络进行训练，可调整神经元中的权重来为各种训练模式产生期望的输出模式（Hebbs，1949）。这通常是通过向系统提供实例来完成的。到达一个正确的结论时，就增大对决策有贡献的神经元的输入的权重，反之亦然。这个反馈系统最终会得到对应于被识别目标的描述的路径。现实中用到了调整权重的各种方法，有些方法前向通过网络，而有些方法反向通过网络（Rumelhart et al.，1986）。

原理上，沿每条路径的各个单元的权重组合描述了高维空间中（对应于输入数量）该类空间的形状。实际中，解释权重很困难，尤其是在系统的内层。处理重叠类（同一组输入可能导致不止一个输出的可能性）的方式与图 12.12 所示的情形完全类似，它报告了每个结果的概率。

许多系统不能解释它们所传递的答案的原因；事实上，很难精确地找到系统中“知识”的位置（就像我们自己的大脑，它似乎广泛地分布在整个网络中）。很难将概率或置信水平赋给各个结果（基于模糊逻辑的情形除外），但神经网络系统通常会体面地失败。神经网络的测试，通常是通过稍微扰动（单个或组合）输入值并观察输出是否仍然稳定来实现的。与经典的专家系统不同，随着可用的信息越来越多且更好地分配了权重，神经系统的求解时间会减少。神经网络方法的主要缺点是，它需要不同寻常的计算机体系结构才能高效地求解。尽管神经网络理论上是并行设备，但许多实际系统使用传统串行计算机通过仿真并行结构来执行所有的计算，因此丧失了速度优势。

稳定神经网络权重通常需要很长的时间，换言之，需要进行大量的训练才能生成可在多数时间给出正确答案的网络。就像上面给出的统计分类方法的要求那样，找到有代表性的训练群体也有相同的要求。训练的质量通常要比如何求解更重要。神经网络与任何一种统计方法之间的唯一不同是，参数空间可分解为各个类区域的复杂性。神经网络可以拟合非常复杂的函数来指定分离区域的表面，在某些情形下它可能比简单的长方体、椭球体或多面体的效果更好，且不会遇到需要无限存储空间、对异常值敏感、难以处理低百分比类以及 *k*NN 方法固有的低速等问题。

上面的描述无疑假设了神经网络的输入值是图像（或图像内各个特征）上的一组测量值，其中每个输入值都是使用前几章中给出的常规步骤得到的。也可使用整幅图像本身作为输入，即使用各个像素值作为输入（Egmont-Petersen et al.，2002；Bollman，2004；France，2004）。此时，网络的最初几层还负责确定应该测量图像中的什么以及如何来测量。这种方法隐藏了来自用户的细节，这看起来很方便，但实际中会使得训练过程更长，而得到的决策难以解读、解释和调试。与涉及离散特征的应用相比，涉及全局颜色和纹理的应用更适合于使用整幅图像作为输入。

12.9 语法模型

还有识别特征的另一种方法，它完全不同于上面描述的参数空间方法（Fu，1974，1982；Pavlidis，1977；Tou & Gonzalez，1981；Schalkoff，1991）。语法模型主要处理通常由特征轮廓或骨架导出的形状信息。形状信息分解成重要的几个片块，并记录这些片块彼此相对的位置。这种记录可类比如下：使用一组字母拼出一个单词，然后根据缺失的、替代的或重排的字母，与其他单词进行比较。

例如，图 12.40 显示了字母“A”的关键元素。由两侧的两个分支点、底部的两个端点及顶部的一个角点，它看起来是字母 A 的形状；字母表中没有其他的字母具有相同的关键点，也没有相同的排列。图中的“弹簧”表明关键点的移动范围很大，但只要它们的相对位置以相同的拓扑顺序保持不变，就不会改变其对识别的有用性。提取这些关键点和它们的排列方式，就可得到该字母的一个非常好的标识符，它对大小和字体相对不敏感。将印刷文字转换为计算机可读文件的一些 OCR（光学字符识别）程序使用了类似的方法。

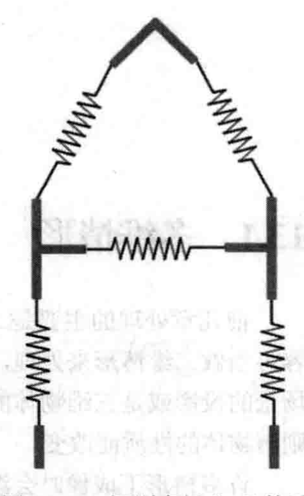


图 12.40　识别字母“A”的重要端点、分支和角点。连接它们的弹簧表明，只要其相对顺序和连接保持不变，移动它们的位置就不影响识别

关键点（如分支和端点）可以直接由骨架得到，如第 8 章所示。角点可以用几种方法分开。图 12.41 中给出了一种方法，其中使用了一个简单的 3×3 核进行卷积，可在任何指定方向选取一个角点。更通用的一种角点查找方法则使用链码横穿特征的骨架或轮廓，这已在第 8 章中介绍过。链码中沿链的某段距离内，给出 90° 方向变化的一个链接序列，被解释为一个角点。使用函数（如多边形或样条）拟合边缘点也可找到角点。

图 12.42 显示了字母 A 到 E 的另一个例子，图中这些字母以多种字体显示，并标示了每一个关键点（端点、分支和拐角）。因此识别每个字符就有了更多的信息；有些点可以去掉（如手写字母引起的极端形状变化）而不影响可读性。开发计算机程序来读取手写字母时，也可使用类似的方案。染色体组型分型中，常常会使用人类染色体中形状和密度模式的语法描述。

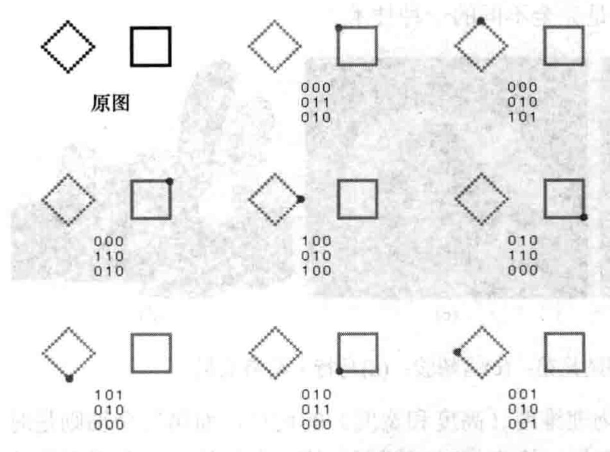


图 12.41　使用一个 3×3 卷积核从骨架或轮廓中提取的角点。每个核找到方形和菱形中标为红色的那些点。一般来说，由核找到的任何图案都具有与期望图案相同的形状

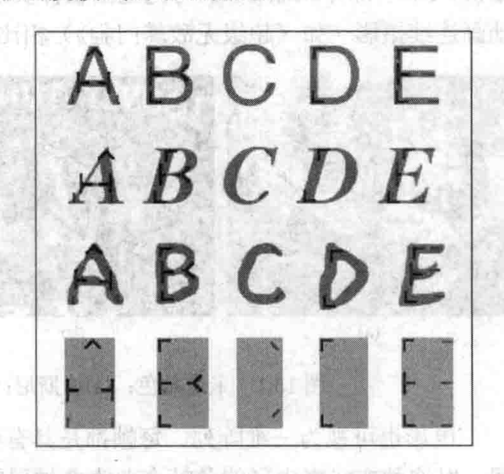


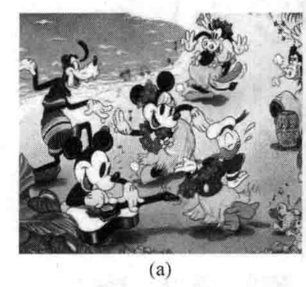
图 12.42　用关键的分支（红色）、角点（蓝色）和端点（绿色）表示的字母 A 到 E

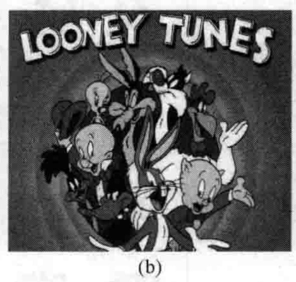
第13章 层析成像

13.1 多维情形

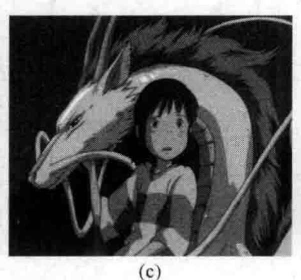
前几章处理的主要是二维图像。图像的主体同样也是二维的，如页面上的印刷文本。有些图像内容可当做二维情形来处理，如基板上散布的颗粒或从卫星上看到的表面。当然，许多二维图像是三维场景的投影或是三维物体的截面。在所有情形下，都可处理和测量像素的二维阵列。测量结果的解释则随物体的性质而改变。

许多情形下成像时会涉及两个以上的维度。例子之一是为经典卡通动画（见图 13.1）使用透明的赛璐珞片。就如迪斯尼和华纳兄弟公司，以及最近宫崎骏［其作品包括《千与千寻》、《悬崖上的金鱼姬》］和马特・葛洛安尼［代表作为《辛普森一家》］的实践那样，卡通技术使用了透明的赛璐珞片［玻璃纸的简称，一种曾经使用的材料，现在可能是计算机图形学程序中的图层］，在这些赛璐珞片上描绘了各种角色和场景元素。有些赛璐珞片（如背景）对整个场景保持固定。其他赛璐珞片则从一帧轻微地移至下一帧，以便移动一个场景元素。通过在各自的赛璐珞片上画上角色（有时会为身体、手、脚等使用单独的赛璐珞片），可降低在两帧间重新画图的工作量。对组成电影各个帧的一堆赛璐珞片进行拍摄，就可形成二维图像。与完全由软件实现的三维渲染（如皮克斯工作室的《玩具总动员》）或黏土动画连续摄影（如《超级无敌掌门狗》）相比，这是完全不同的一种技术。


(a)



(b)


(c)


(d)

图 13.1 卡通角色：(a)迪斯尼；(b)华纳兄弟；(c)宫崎骏；(d)马特・葛洛安尼

电影也可视为三维阵列。每帧都是具有两个物理维度（高度和宽度）的切片，而第三个轴则是时间。以这种方式将电影的各帧合并为叠加层的优势是，检查与时间轴平行的一个切片，就可很容易地看出场景的过渡。此外，对于压缩目的而言，与动态场景相比，静态或缓慢移动的部分场景可以单独压缩，进而大大降低文件尺寸。

重叠图像的各层并不是三维的，也不由单独的像素阵列组成，更不是体素的单个集合。每层可以包含完全不同的信息。例如，地理信息系统（GIS）可包含地图、卫星图像（包括红外波长）、调查记录和纳税记录、服务区的位置和特定特征的位置等。谷歌地图就是一个我们熟知的例子（见图 13.2），它可根据消费者的需求，包含更多的此类数据。它提供的街景服务可基于底层地图的定位，在任何道路上的任何位置显示街景的全方位视图。

卫星图像通常会以所选的几个波长来记录图像。图 13.3 显示了新墨西哥州洛斯・阿拉莫斯的几幅图像（它们由 Landsat 7 卫星上的增强型专题绘图器拍摄）。将通道 1（蓝色）、通道 2（绿色）、通道 3（红色）合并为 RGB 彩色通道后，得到的真彩色图像中清晰地显示了森林火灾的烟羽。在通道 2、通

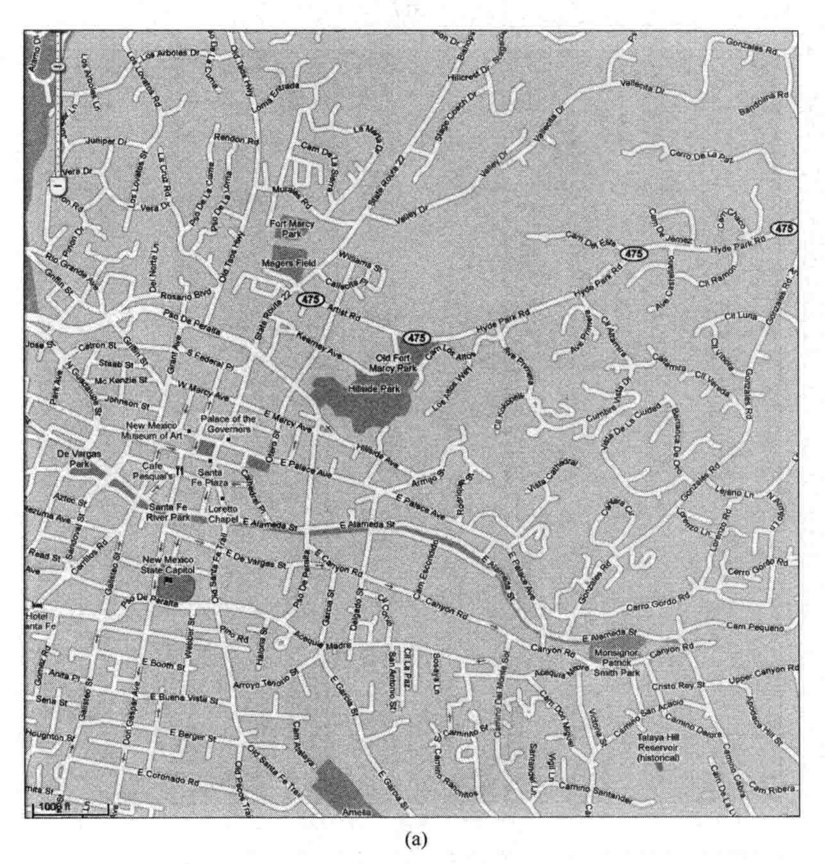
道 4（近红外）和通道 7（中红外）合并得到的伪彩色图像中，可看到森林火灾多发区中的热点和燃烧痕迹。

以这种方式来组合多幅单独的图像（要求这些图像是对齐的），是一种有效可视化的强大工具，但它不会将这些图像真正转换为三维数据集。而要得到三维数据集，则需要在图像中的每个像素位置得到光谱强度与波长的完整关系（Bonnet et al.，1992；Kim et al.，2003；Bearman，2003；Booth et al.，2010）。技术上使用扫描分光光度计可以实现这一目的，有时显微镜检查和取证学方面会用到这种技术，但会相当耗时，且得到的数据文件非常庞大。

尽管这是包含有两个以上维度的一个数据集的例子，但如果一个维度不是测量的距离或时间，那么很少会有能对其进行处理或分析的工具。对一些特定波长的亮度空间分布成图，可生成二维图像。对许多目的而言，选取感兴趣波长并直接收集对应的二维图像更为有效。完整数据集的优点是，日后需要时可以随时提取这些信息。

电子探针、扫描电子显微镜和分析透射电子显微镜完全采用这种工作方式。尽管它们能在任何点进行详细的分析，且原理上可以在扫描图像中的每个点处进行分析并记录数据，但最常见的操作模式是选取一些对应于感兴趣元素的波段，并将亮度记录为位置的函数，生成如第 1 章中图 1.32 所示的“X 射线图”。其他扫描分析仪器也可以这样操作，譬如在透射电子显微镜和一些核医学成像技术中记录能量损失谱。

在本章和接下来的几章中，三维成像主要关心三个空间维度。本章介绍使用层析技术的真正三维体素成像。第 14 章介绍一组平面图像的可视化和用途。第 15 章介绍表面成像和测量，其中高程值的二维阵列可视为“二维半”图像。



(a)

图 13.2　新墨西哥州圣达菲地区的谷歌地图显示：(a)带有标记的道路图

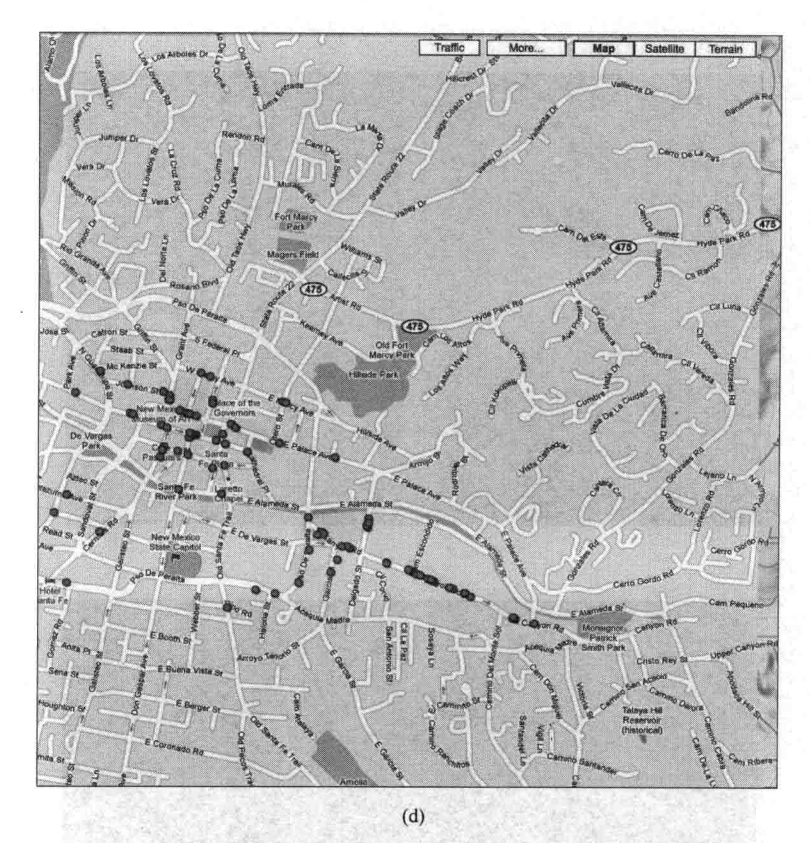


(b)

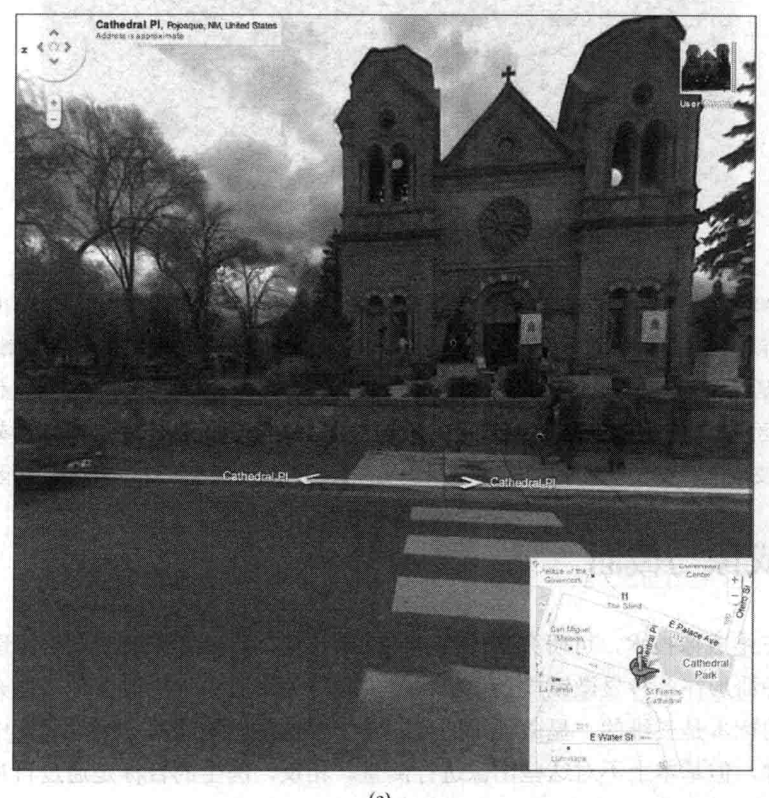


(c)

图 13.2　新墨西哥州圣达菲地区的谷歌地图显示：(b)卫星图；(c)地形图（续）



(d)



(e)

图 13.2　新墨西哥州圣达菲地区的谷歌地图显示：(d)定位艺术画廊的标记；(e)大教堂街景（续）

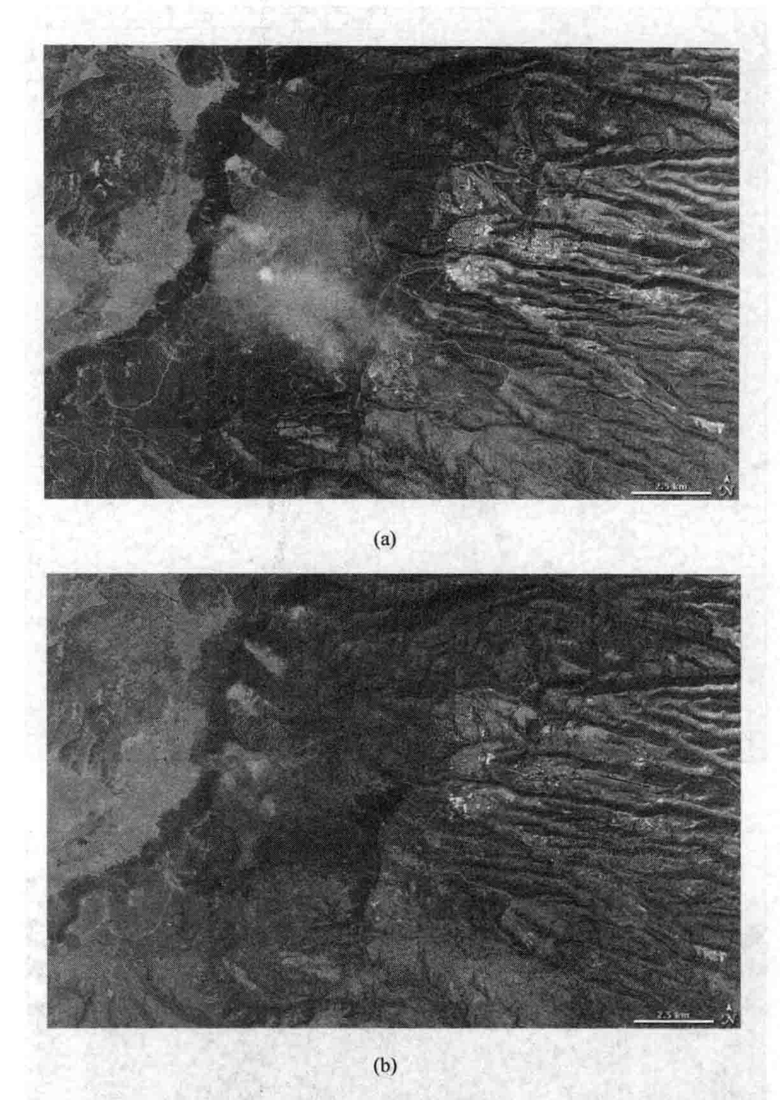


图 13.3 2000 年 5 月 9 日新墨西哥州洛斯·阿拉莫斯的几幅卫星图像中显示了 Cerro Grande 火灾的情形（这几幅图像由美国航空航天局和美国地质调查局的陆地卫星专题制图仪拍摄）：(a)通道 1、2 和 3（可见光，直彩色）；(b)通道 2、4、7（1 可见光、2 红外，伪彩色）。森林区为暗绿色，草本植被为浅绿色，牧场为粉红色或浅橙色，发达的城市区域为蓝色、白色或紫色，高尔夫球场为亮绿色。最近燃烧过的区域为黑色，红色表示火势最强烈的区域

13.2 体积成像与截面

使用二维或三维成像技术，可研究固体物体的三维结构。前几章中的有些二维图像就是通过三维结构的截面。这种情形在各种显微镜检查中很常见，此时为形成二维图像会使用抛光的平面或切割的薄片。这样采样的标本是三维的，显微学家的目的是了解这些三维结构。医学成像中也广泛使用切片和投影的二维图像，但基本上不对这些图像进行测量。相反，医生的目标是通过目视来定量检查器官和骨骼偏离正常情形的程度，此时会严格控制图像的方向，以便使得比较更为容易。

存在强大且实用的定量工具，它们可有效地解释二维图像的测量结果，进而表征三维结构。如第 9 章中总结的那样，这些工具可以测量结构的体积比、相交面的面积、薄膜的平均厚度，甚至仅在某个随机截面中出现的颗粒的粒度分布。但结构的许多方面（定量的和定性的）不能由二维截面图像得到。拓扑性质就是这样的一类信息，体积中单独物体的数量这种“简单”的值也是这样的一类信息。单个二维截面中出现的特征并不能揭示实际的三维结构，如图 13.4 所示。

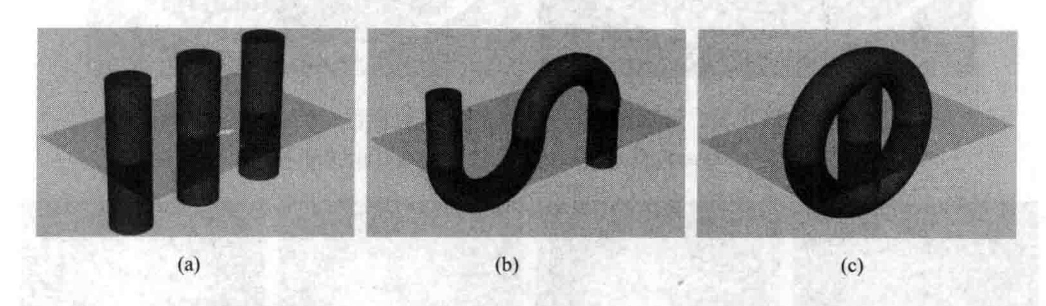


图 13.4　通过三个不同结构的截面显示了相同的二维截面图像：(a)三个分开的物体；(b)简单连通的一个物体；(c)多连通的一个物体

第 9 章显示了将圆形分布展开，进而确定样本中尺寸分布情况和球体总数的一种方法。但该过程要求物体都是球体或已知的某种几何形状。如果不知道特征的形状，且特征的形状是变化的，或特征的形状与特征的大小相关，那么就不能使用这种方法。

此外，像数值测量数据那样，它们不能为普通查看者提供真正意义上的结构。如第 2 章指出的那样，人类绝对是视觉动物，即人们希望看到三维结构。我们看到的世界是三维的，我们习惯于查看外部表面，或偶尔查看透明介质，而不擅长于查看薄的截面或抛光的截面。假设你正站在某个繁忙的街道上，用一个假想的平面横切交通流，此时脑中会存在这样一个截面，但截面中什么也没有。当人、汽车和自行车通过该平面时，它们看上去像什么？

如果你有很好的几何感，或具有放射科医师或绘图员的经验，那么当该平面穿过人体躯干、汽车发动机时，你可能会准确地设想该截面的外观。但多数人很难想像这样的一组截面，且在被要求画出截面中的简单形状时，结果会非常不准确。

如果你对此存疑，可对自己进行简单的测试。找到一些脆谷乐麦圈或螺旋形面条，或形状规则的其他食品（见图 13.5）。然后与一种不透明的基质（如软糖）混合并搅拌。当混合物变硬后，试着画出穿过混合物的随机切片上的内容。混合物硬化后，切开样本，并将实际的截面与你所画的内容进行比较。我与学生进行这一实验时，学生们都认为通过物体的截面是平行的且垂直于主轴，并通过物体的几何中心。对于脆谷乐麦圈而言，几乎没有由两个并排的圆或一个环组成的实际截面，且截面甚至不是凸状的（见图 13.6）。至于螺旋形面条，多数人都未意识到通过弯曲表面的截面会产生直线。

和这一特殊视觉任务一样困难的是，很少有人能反向思考这一过程：给定一组随机截面的剖面，在脑中重建三维结构的正确表示。即使是那些对二维图像本身得心应手的人，也无法很好地完成这一过程。在二维图像中，识别和理解并不需要关于三维世界的单独知识库。观察抛光金属样品中的树突外观，或电子显微镜截面中细胞线粒体的典型外观，或医学图像中的内脏（此类图像几乎总是轴向、冠状或矢状图像，如图 13.7 所示），并不一定意味着会熟悉这些物体的复杂三维形状。由二维截面图像所形成的三维解释，存在大量出现错误的示例。

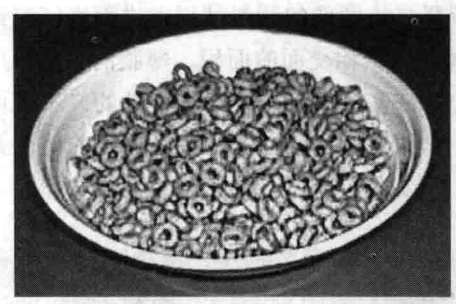
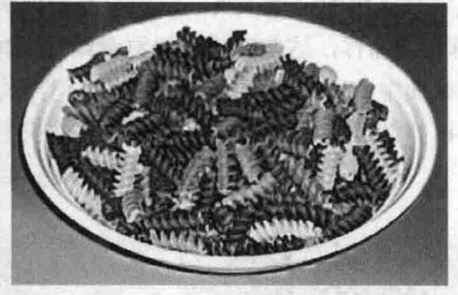



图 13.5　三维形状和二维截面图像关系实验所用的食物（脆谷乐麦圈和螺旋形面条）。与多数自然结构相比，这两种食物的形状更简单、更均匀

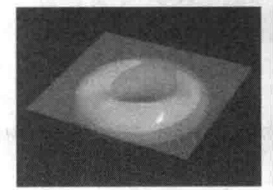
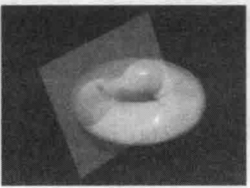
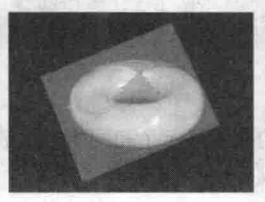
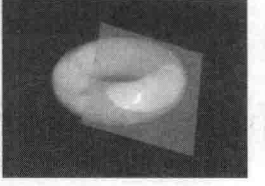
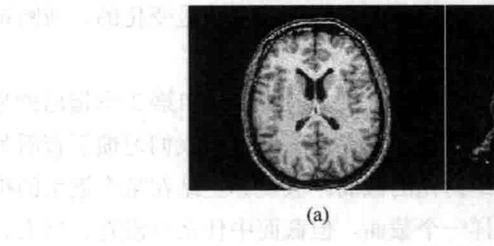
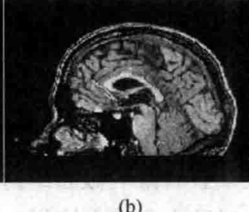
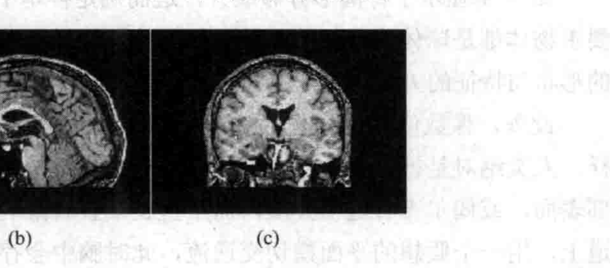





图 13.6　穿过圆环体的几个截面，产生了一个或两个凸状或凹状相交面

(a)　(b)　(c)

图 13.7　查看医疗切片图像的标准方向：(a)轴向；(b)矢状；(c)冠状

由于使用二维图像来研究三维结构的这种困难性，因此我们希望能够进行三维成像。如本章所述，三维成像可通过一次性收集所有三维信息来直接实现；或如下一章中所述，通过收集一组二维（切片）图像并组合它们来间接实现。有时这些切片必须通过物理手段切割才能得到（如考古学家挖掘现场），但有时可以是非破坏性的（如使用共焦光学显微镜或核磁共振成像仪）。获取二维图像来收集三维成像所需数据的方法还有很多，将这些信息呈现给用户的方法也有很多。第 14 章中介绍了这些方法。大量应用表明，无论是对于多数观察者还是对于多数应用，并不存在最好的方法。

重建一个（由一系列投影数学重建的）物体内的内部结构信息的任何方法，通常称为断层成像。使用相同的基本方法可得到真正的三维体素阵列（二维图像中像素的三维类似物），或由一系列一维线投影得到二维图像。后一种方法常用于医学成像，它是目前为止最常见的 CT（计算机断层扫描）应用。但同样的技术也适用于三维成像和各种成像信号。

医学断层扫描主要使用 X 射线吸收、磁共振（MRI）、光子（SPECT）或正电子（PET）发射和声波（超声波）。在最新出版的教材中，我们可以找到各种方法的介绍（Jan，2005；Rangayyan，2005）。其他领域的应用和研究会使用电磁辐射的不同频率，从 X 射线和伽马射线（纳米波长），到可见光甚至微波辐射（厘米以上波长）。除光子外，断层扫描常使用电子和中子。除粒子吸收或辐射外，断层扫描也可基于辐射的散射或发射。混合流体的断层扫描有时可通过测量安全壳周围多个点间的电阻或阻抗来进行（Mann et al.，2001；York，2001）。

通过人为爆炸或人为振动来产生声波，可以对地层成像，天然地震这种噪声源也可用于进行地震

层析成像，即对地下的断层、火成岩的密度成像，并找到地幔和地核间的不连续性。使用监测月球地震和火星地震的设备，可揭示它们的内部结构，而对根据太阳震动来了解其结构的研究一直在进行。

13.3 重建的基础

X 射线吸收断层成像是最古老、最广泛使用的方法之一，因此这里使用它来说明各种参数、伪影和性能。由计算机辅助断层成像或（使用磁共振、声波、同位素发射、X 射线散射或电子束的）类似方法产生的图像值得特别注意。这些图像的具体形成方式如下：对沿线的投影进行测量，并使用计算机处理各个测量信息，这种穿透人体的方式是非破坏性的，但必须展开才能看到内部结构。下面介绍医学应用和工业测试中所用 X 射线吸收断层成像的数学原理（Herman，1980；Kak & Slaney，1988，2001；Natterer，2001；Natterer & Wubbeling，2001；Hsieh，2003，2009）。类似的方程组和求解方法适用于其他信号模式。

吸收断层成像基于辐射或粒子径直穿透样本后强度下降的物理过程。在某些类型的断层成像中，路径并不是直线，且重建是沿曲线（如磁共振成像和 X 线散射断层成像）或同时沿多条直线（如地震层析成像）进行的。这就使得方程和图形更为复杂，但基本原理相同。

X 射线穿透材料，并根据所遇到的成分和密度沿路径吸收。强度（光子数/秒）以线性衰减系数 μ 下降，对于感兴趣标本，这个系数并不一致，而是有些空间变化，因此对于三维结果它可写为 $\mu(x,y,z)$，或对于通过物体的二维平面可写为 $\mu(x,y)$。线性衰减系数是密度和质量吸收系数的乘积，它取决于局部元素组成。在医学断层成像中，这种组成的变化很小，主要由密度变化来生成图像。对于工业应用，组成的变化很大（甚至包含空隙）。沿这一分布的一条直线路径所测量的强度为

$$\int \mu(x,y)\mathrm{d}S = \ln \frac{I_o}{I_d} \tag{13.1}$$

式中，I_o 是入射强度（来自 X 射线管或放射性同位素），它是已知的且通常保持不变，I_d 是检测到的强度。上式称为射线积分方程，它描述了沿一条穿过物体的线投影的结果。

如果测量了一系列平行直行（通过扫描发射源和检测器一次测量一条直线，或使用许多检测器测量所有直线），就可得到称为视图的强度剖面。如图 13.8 所示，该函数通常画为强度的倒数，或沿每条直线的吸收和。函数写为 $P(\varphi, t)$ 表明，当射线到达物体的不同位置时，函数会沿方向 t 随位置的不同而变化，射线源围绕物体以角 φ 旋转或将物体旋转角 φ，可从不同的方向来查看图像。

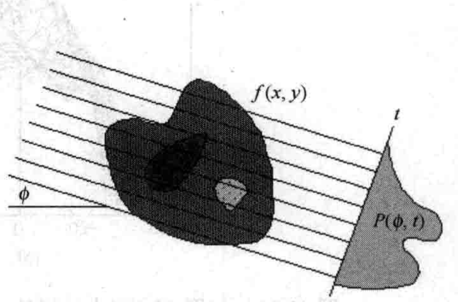


图 13.8 以视角 φ 穿过物体的一组投影响，形成了密度函数 P

每个视图都是所测衰减的一维剖面，而衰减与某个特殊角度的位置有关。多个这样的视图集合就可表示为一幅二维图形或图像，其中的一个轴是位置 t，而另一个轴是角度 φ。这种图像称为二维切片的正弦图或 Radon 变换。图 13.9 和图 13.10 显示了一个简单的例子。如图 13.9(a)所示，构建的平面图形称为体模，它用于评价重要的变量和由投影信息重建物体切片的不同方法。这样的体模通常可模仿我们真正感兴趣的重要结构，有助于我们评估重建算法。请将它与脑瘤的真实 MRI 图像即图 13.9(b)进行比较。

图 13.10(a)中所示的各个投影剖面显示了角度改变时的一些变化，但这种现象很难解释。图 13.10(b)中的正弦图组织了数据，以便可更容易地进行检查。正弦图一词源于经过体模内不同结构的投影位置的正弦曲线变化，它是旋转的函数，在该例它非常明显。Radon 变换源于这种成像方法的原理是由 Johann

Radon 于 1917 年给出的。但他提出的方程并不提供实现重建的实用途径，因为这些方程要求连续的投影阵列，直到 Hounsfield 和 Cormack 开发了一种实用的重建算法和硬件后，CAT 扫描才成为常规的医疗手段。A. M. Cormack 在 1963 年至 1964 年于塔夫茨大学提出了数学上可管理的重建方法，而 G. N. Hounsfield 在 1972 年于英国的 EMI 公司设计了一台有效的仪器。他们共同于 1979 年获得诺贝尔奖。

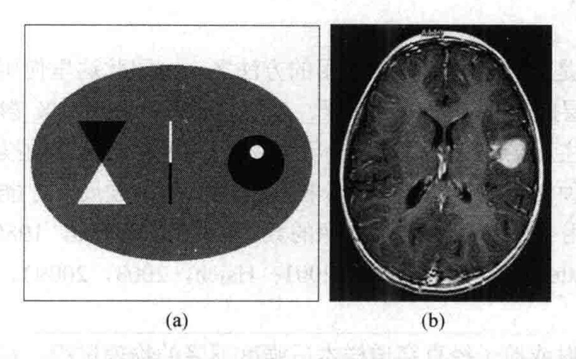


图 13.9 (a)体模或已知密度几何形状的测试物体；(b)头部截面中的肿瘤（亮区）

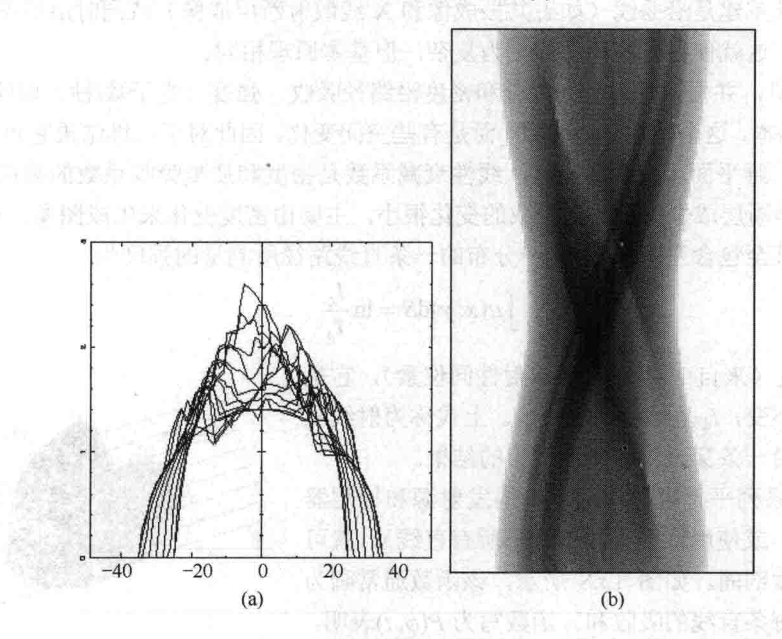


图 13.10 (a)图 13.9(a)中体模的 16 个衰减剖面；画出 180 个这种剖面（每个剖面作为一条水平直线）得到的正弦图或 Radon 变换

在某个观察方向，一组投影数据的傅里叶变换可写为

$$S(\phi,\varpi)=\int P(\phi,t)\mathrm{e}^{-\mathrm{j}2\pi\omega t}\mathrm{d}t \tag{13.2}$$

Radon 变换证明它也可以写为

$$S(\phi,\varpi)=\iint f(x,y)\mathrm{e}^{-\mathrm{j}2\pi\omega(x\cos\phi+y\sin\phi)}\mathrm{d}x\mathrm{d}y \tag{13.3}$$

它恰好是函数 $f(x, y)$的二维傅里叶变换 $F(u, v)$，约束条件为 $u=\omega\cos\varphi$ 和 $v=\omega\sin\varphi$。因此，这就是投影的直线方程。

这一关系表明，从体模的原始图像开始，按第 6 章中给出的方法形成其二维傅里叶变换，然后沿从原始点到方向 φ 的法线这一径向方向观察该图像中的信息，可得到函数 S，它正好是实空间中 φ 方向上的投影数据的一维傅里叶变换。实际使用方法是，在许多方向测量投影 P，计算出其一维变换 S，

将 S 的复系数画到相应方向上的一幅二维变换图像上，并在测量足够的方向后，执行二维傅里叶反变换，来恢复切片的空间域图像。这样就可由投影的数据来重建切片图像，从而得到一幅非破坏性的内部图像。这就是层析成像的原理。

图 13.11 举例说明了傅里叶变换数据和重建之间的关系。以相同角度间隔拍摄了 8 个视图或一组投影，计算出每个投影的傅里叶变换并画到一幅二维复图像中，然后即可完成重建。因为视图数量有限，故图像质量一般。使用 1° 间隔的 180 个视图时，结果相当不错。因由频率空间图像中存在间隙，因此仍出现了伪影。在高频处（远离频率空间图像中心的位置），这种丢失的信息尤为明显，因为此处各个视图的直线间距更宽。所有层析成像的重建过程都对视图数量敏感。

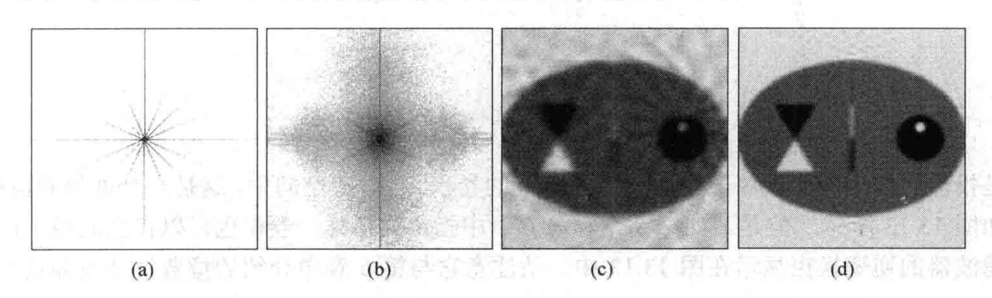


图 13.11　使用图 13.10 中的投影数据在频率空间中重建。将不同角度的一组投影或视图的一维(复)傅里叶变换画在一幅二维频率域图像上，即可完成重建：(a)频率空间中的 8 个视图；(b)频率空间中的 180 个视图；(c)由图(a)得到的重建结果；(d)由图(b)得到的重建结果

通过收集足够多的视图并执行这一傅里叶空间重建，就可进行层析成像。但在实际中，很少有系统采用这种工作方式。需要更小计算量且更易于“实时”处理的一种数学上等效的过程，称为滤波反投影。这是绝大多数医学扫描仪和有些工业应用中采用的方法。

反投影的原理很简单。在视图的每个投影中画出的衰减，是样本结构沿各条直线或射线的积分。仅由一个投影就可知道直线上衰减出现的位置是不可能的，但可以沿直线均匀地分布测量的衰减。如果仅为单个视图这样做，那么所得结果的意义不大。但如果对来自几个视图的投影这样做，那么叠加的密度或衰减值应对应于结构中出现的各个特征。

图 13.12 给出了同一个体模的这一结果。在投影中可看到中空（亮）的致密（暗）圆柱。叠加来自几个视图的数据，可在重建图像中画出该圆柱。但该结果有一个问题。原始体模中均匀区域的衰减或密度并不恒定，而是越靠近截面的中心，衰减或密度会变得越大。同时，边缘会被模糊。

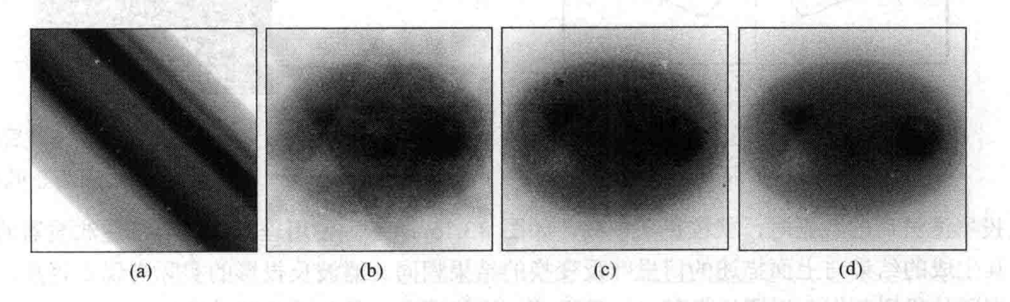


图 13.12　空间域中的反投影。每个视图中的衰减值沿每条投影直线投影回物体空间。将来自多个观测方向的数据加在一起，确实显示了几个主要特征，但图像却是模糊的：(a)1 个视图；(b)6 个视图；(c)30 个视图；(d)180 个视图

该问题的起因可用几种不同但等效的方式来描述。来自所有视图的投影主要作用于图像的中心，因此图像中心处所有的投影会重叠在一起。使用散焦光学系统（其模糊或点扩散函数与 $1/r$ 成正比，其

中 r 是频率，或到频率变换中心的距离）来观测图像时，效果相同。

第 6 章介绍了如何从图像中消除一种已知的模糊：对频率空间变换应用一个反函数来校正该模糊，此时要在再次变换前衰减低频。基于傅里叶的方法，并根据极坐标写出反变换有

$$f(x,y)=\int_{0}^{\pi}\int_{-\infty}^{+\infty}S(\phi,\varpi)\cdot|\varpi|\cdot e^{j2\pi\omega t}\,d\varpi\,d\phi \tag{13.4}$$

或根据 x 和 y 写出反变换有

$$f(x,y)=\int_{0}^{\pi}Q_{\phi}(x\cdot\cos\phi+y\cdot\sin\phi)\,d\phi \tag{13.5}$$

式中，

$$Q_{\phi}=\int_{-\infty}^{+\infty}S(\phi,\varpi)\cdot|\varpi|\cdot e^{j2\pi\omega t}\,d\varpi \tag{13.6}$$

这正好是投影的傅里叶变换 S 与频率的绝对值$|\omega|$的卷积。在频率空间中，这是一个理想的逆滤波器，其形状如图 13.13 所示。但是，如第 5 章和第 6 章中指出的那样，卷积也可以在空间域中应用。这个理想滤波器的逆变换也显示在图 13.13 中。请注意它与第 5 章中介绍的拉普拉斯或高斯差的形状类似。

就像一个一维的核或权重集一样，该函数可以乘以投影 P，就如将核应用到第 4 章和第 5 章中的二维图像中那样。权重乘以这些值，然后将总和保存为滤波后的投影中的一个点。对投影集或视图中的每条直线重复该过程。图 13.14 显示了投影数据的结果，该结果以正弦图的形式出现。边缘（高频）得到了强烈增强，而低频信息则被抑制。

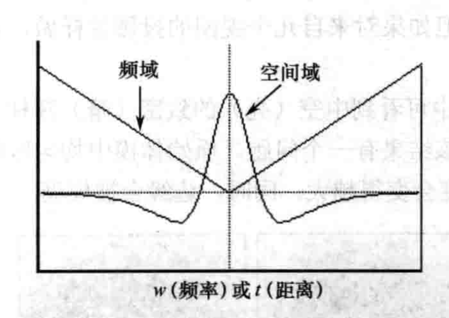


图 13.13　一个理想的逆滤波器（它选择性地删除低频）及其空间域的等效核

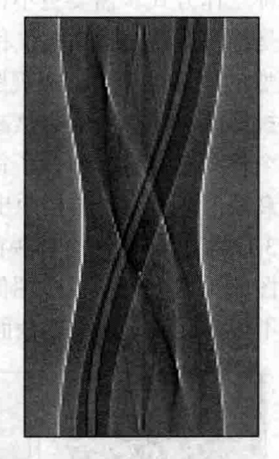


图 13.14　来自图 13.10(b)的滤波后的投影数据，显示为正弦图

反投影滤波后的数据时，就校正了模糊，如图 13.15 所示。使用理想滤波器或逆滤波器的滤波反投影，其生成的结果与上面描述的傅里叶反变换的结果相同。滤波反投影的实际实现更容易，因为来自每个视图的投影数据可以通过卷积（一维运算）进行滤波，且随着视图被获取，这些数据会传回整个图像，而无须存储傅里叶方法或之后的反变换所需要的复（即实值和虚值）频率空间图像。

注意，在图 13.15 中，图像的质量和视图的数量对伪影的影响，与图 13.11 中频率空间方法对其的影响相同。在数据中缺少噪声和下面讨论的其他影响情形下，这两种方法是完全等价的。滤波反投影法是当前多数 CT 仪器所采用的方法。

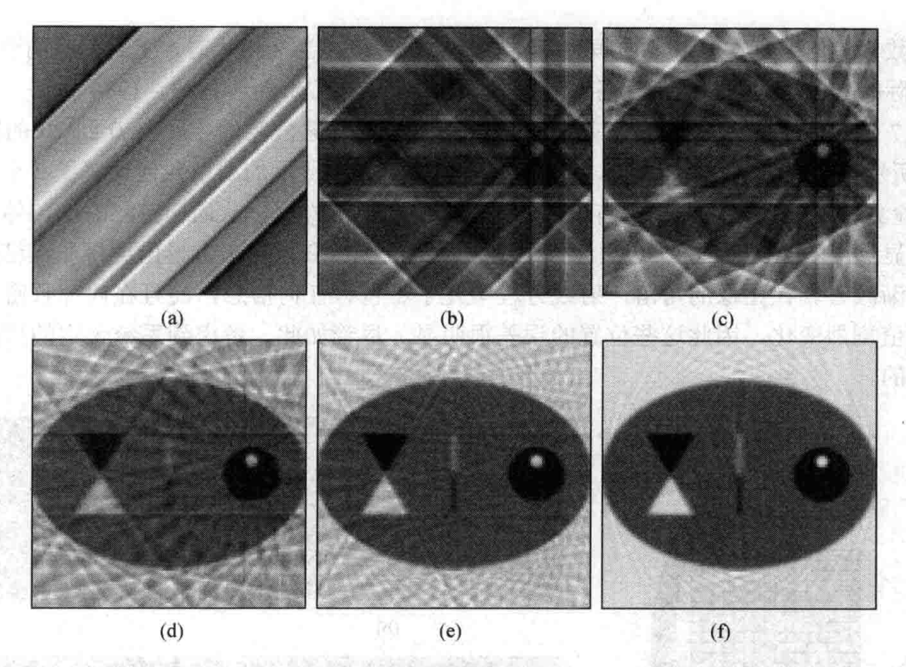


图 13.15　滤波后的反投影。该方法与图 13.12 中的方法相同，只是每个视图的值已用图 13.13 中的函数通过卷积滤除：(a)1 个视图；(b)4 个视图；(c)8 个视图；(d)16 个视图；(e)32 个视图；(f)180 个视图

13.4　代数重建方法

求解图像中每个位置的密度（更精确地说，求解线性衰减系数）的问题，也可视为求解一个联立方程组。每条射线的积分（在此处考虑的有限情形下的求和）提供一个方程。沿射线的体素的衰减系数之和（每个系数乘以一个加权因子，加权因子的大小与射线穿过该体素的实际路径长度有关），等于所测量的吸收。图 13.16 说明了体素和射线积分方程之间的关系。

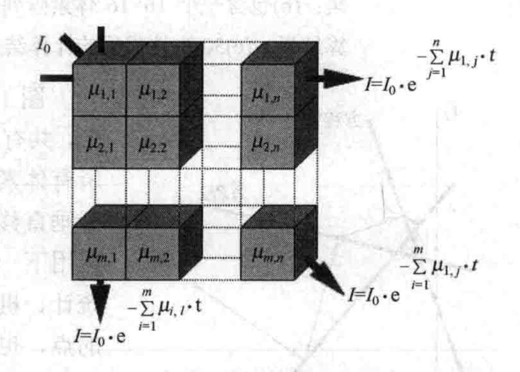


图 13.16　在标本的一个平面截面中的体素的示意图以及对沿着通过阵列的几条路径衰减求和的射线积分方程

这个方程组中的未知数的数量，是通过标本的切片中的体素的数量。方程的数量是射线积分的数量，它通常是每个投影剖面的检测器数量与视角数量的积。这是方程数非常多的方程组，但许多权重为零（任何特殊的射线积分方程中并不涉及多数体素）。方程的数量与未知数的数量很少相同，但所幸的是，当这样的稀疏方程组欠定或超定时，可使用许多实用且经过测试的计算机方法来对它们求解。

本书的目的并不是比较各种求解方法。使用代数重建技术或 ART（Gordon，1974）这种最简单的方法，可充分理解这种求解方法。在这种方法中，方程可以迭代求解。方程组可以写成

$$\boldsymbol{A}^{m\cdot n}x^{n}=b^{m} \tag{13.7}$$

式中，n 是体素的数量，m 是投影的数量，$\boldsymbol{A}$ 是每个体素对每条射线路径的贡献的权重矩阵（可为任何特定的仪器和几何预先计算该矩阵）。体素值是 x 值，而投影测量值是 b 值。经典的 ART 方法由前一组 x 值将每个 x 值的迭代集合计算为

$$x^{k+1}=x^{k}+\boldsymbol{A}_i(b_i-\boldsymbol{A}_i^{\lambda}x^{k})\|\boldsymbol{A}_i\|^2 \tag{13.8}$$

式中松弛系数 λ 的值通常在 0 和 2 之间，用于控制收敛速度。当 λ 很小时，上式就等同于传统的最小二乘解。实际的考虑因素（包括应用各个方程的顺序）详见（Censor，1983，1984）。

图 13.17 显示了该方法的一个简单的例子。对 16×16 体素阵列给定了从 0 到 20 的密度值，如图 13.17(b)所示，在 0°、90° 和 180° 视角，为图 13.17(a)中所示的扇束几何计算了 3 个投影集。对于有 25 个检测器的阵列，这给出了有 256 个未知数的 75 个方程。从最初猜测的均匀体素（密度为 10）开始，显示了 1 次、5 次和 50 次迭代的结果。空洞区域和内部正方形出现得非常迅速，边界的清晰度也逐渐改善。在图像的角落，射线方程中几乎不包含任何信息，而且在内部致密正方形的角点处，衰减值剧烈变化，因此这些位置的误差很明显。尽管如此，考虑到系统欠定的程度，结果仍是令人鼓舞的。

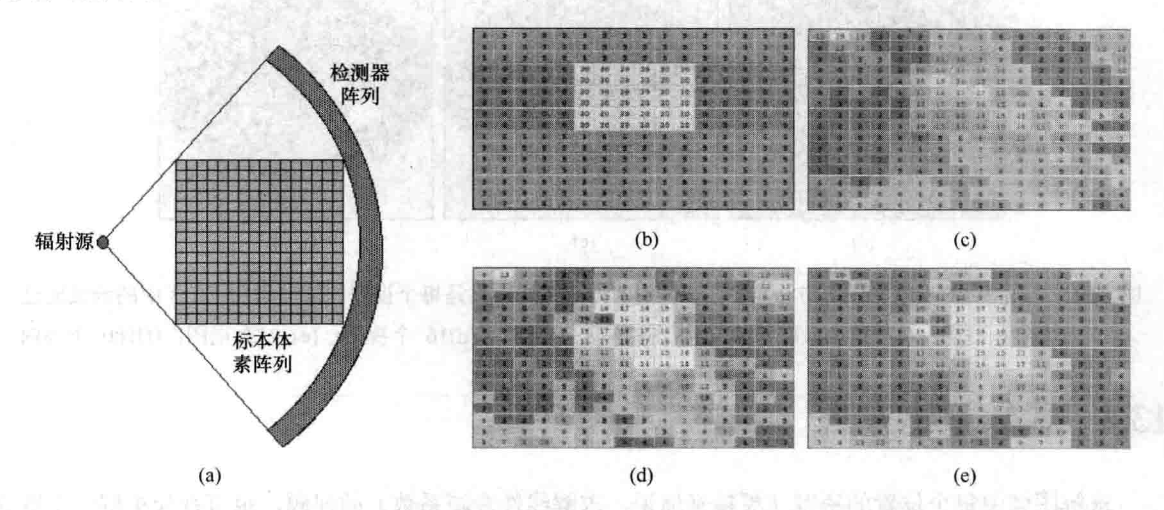


图 13.17　迭代求解的应用举例。(a)在 0°、90° 和 180° 视角，为 25 个检测器的阵列计算了三个投影集；(b)包含一个 16×16 体素阵列的模拟标本；(c)1 次迭代后的计算结果；(d)5 次迭代后的计算结果；(e)50 次迭代后的计算结果

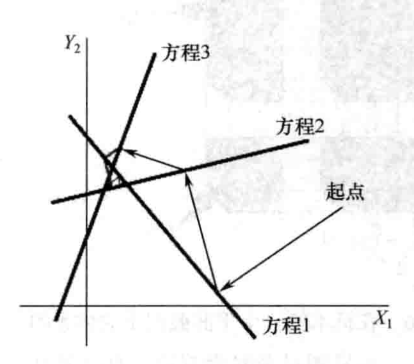


图 13.18　迭代求解一组方程的 Kacmarz 方法示意图

图 13.18 中说明了此类解的 Kacmarz 方法，对于这种适度的情形，共有三个方程和两个未知数，且 λ=1。从一些初始的猜测值（如所有体素具有相同的衰减值）开始，应用其中的一个方程。这相当于垂直移动到表示该方程的直线。然后将这个新点用做一个起点来应用下一个方程，以此类推。在实际情形下，由于各种测量、计数统计、机器变化等的精度有限，因此这些方程并不都满足一个完美的点，也就没有某个点能代表稳定解。相反，解会收敛到一个区域（该区域主要位于各条直线间的区域内）并在该区域中摆动。但在具有某些带噪方程的高维空间中，可让解离开这个区域并在多次迭代后发散。

在具有许多维度的真实情形下，收敛速度可能较慢。使用迭代代数技术的最大困难是决定何时停止迭代。从逻辑上讲，目标是继续下去，直到答案尽可能好，但由于不知道“真相”，因此不可能精确地判断这个停止点。有些方法会在每次迭代后检查计算出的图像中的变化，并试图根据这一变化来判断何时停止（例如，当体素值的归一化全变差下降到低于某个任意限制时，或当它由上一次迭代开始增加时）。有些情形下，这种方法会出现严重的误差，但多数时间执行良好，因此常被人们使用。继续迭代的代价不只是计算成本，还

有对于某些数据集解会发散（离开交叉点附近有界区域）的可能性。

既然代数方法存在诸多弊端，而滤波反投影方法又相对简单和直接，那么为何还要使用代数方法呢？代数方法（如 ART）有一些潜在的优势。首先，滤波反投影法及其体现的傅里叶变换方法，需要视图的数量相当大，且要求这些视图是等间隔的，以便使用数据更好地填充频率空间。物理限制导致的一些角度或整组角度的丢失，会在滤波反投影时出现问题，并会引入明显的伪影。ART 方法仍然可以产生一个可以接受的重建。投影欠采样的重建图像中，有些部分可能会缺少细节，但伪影不会遍布整幅图形。事实上，可接受的重建通常是由少数几个视图得到的（例子如下所示）。

ART 的另一个优点是应用约束条件的能力。例如，由于有限的测量精度，在滤波反投影或傅里叶变换方法中，可为某些体素计算负密度值（衰减）。这样的值没有物理意义。在迭代代数方法中，任何这样的值都可以限制为零。在图 13.18 的示意图中，它将解限制在具有正值的象限中。

此外，可以应用先验知识。如果已知标本中的某些密度和衰减值对应于特定的材料，那么这些值可以很容易地约束到一致。还可以包含任何几何信息，如物体的外部尺寸（例如，将物体边界外部强制为零密度）。

也可建立尺寸和间距并不都相同的体素网格。例如，这种设置可能允许在物体内部使用更小的体素间距来得到更好的细节，而在物体外部使用更粗糙的网格，反之亦然。这仍然允许计算外部材料对射线积分的贡献，但未知数数量的降低，可为给定数量的视图和投影产生更好的解。当需要符合特定目标形状和对称性时，也可使用非立方体素。

代数方法的灵活性及其在工业断层扫描设置中使用先验信息的特殊能力，可弥补其计算缓慢和需要大量计算的不足。体素权重（$\boldsymbol{A}$ 矩阵）的计算可能很烦琐，尤其是对于扇形束或其他复杂的几何形状，但并不比同样情形下的反投影更烦琐，而且它是一次性计算，计算结果可保存起来，并用于那些使用相同几何的重建。使用求解方法而非此处描述的简单迭代方法，可改进解的稳定性和收敛性。

13.5　最大熵

还有其他方式来求解这些巨大的稀疏方程组。其中一种方式称为最大熵方法。第 4 章中曾介绍过最大熵，当时它是消除二维图像中噪声的一种图像工具。若假定存在可用做约束条件的相关先验信息，那么贝叶斯定理就是最大熵方法的基础。在没有先验信息可用且噪声是支配因素的情形下，贝叶斯定理就导致了经典的逼近法或“最小二乘”逼近法。先验信息的使用导致了一种改进的方法。

最大熵方法的哲学思想源于贝叶斯统计和信息论，也源于统计热力学的吉布斯概念（Jaynes，1967）。总之，它可描述如下：找到可用（与已知约束一致的，如光子总数、在任何点处的非负亮度或密度、检测器或测量过程中涉及的物理学等）和具有最可能配置值的结果（图像像素中的亮度分布、体素阵列中密度值的分布等）。

这一概率定义为能够以多数方式形成。对于一幅由光子形成的图像，所有光子都被认为是不能分辨的，且它们到达的顺序不重要，因此可用许多方法将光子分布到各个像素。对于某些亮度模式（图像）而言，与其他亮度模式相比，形成模式的方式的数量要多得多。具有较大多样性的图像，会有较高的熵。可用更多的方式来自然形成它们。熵定义为 $S=-\sum p_i \log p_i$，其中 p_i 是亮度值 i 的像素百分比。

从简单的统计观点来看，最可能的图像是对所有像素获得相同平均数量的光子，产生均匀的灰色场景。但约束条件可能不允许这种结果，如实际记录的所测亮度模式就是这样一个约束条件。计算场景与测量场景之差允许设置一个上限，通常基于检测器的估计噪声特性、光子的数量等。找到具有最

大多样性的可用场景就是最大熵方法。

例如，在用不同视角的射线积分方程求解一个物体的层析重建时，会有一个有许多未知数的较大联立方程组。替代正式求解联立方程组（例如，采用传统的高斯一约当消元法，这种方法步数太多，因此不实用），最大熵方法彻底改造了这一问题。从任何初始猜测值开始（在最“良好”的情形下，这个猜测值的好坏对最终结果影响不大），然后迭代从这一点开始，找到具有较高熵的另一个解（在约束条件定义的可用解类别内）。在由参数（所有体素的值）定义的空间中，决定以哪一种方式来移动通常由拉格朗日乘数来完成，方法是取偏导数，并总是试图“向上”移动，此处用于评估每组值的目标函数就是熵。

通常发现，具有最大可用熵（即约束条件允许的熵）的解，很难对付由这些约束条件形成的一个或多个边界，且如果放宽约束条件，这些解便会向更高处移动（趋于一幅更一致的图像）。知道或假设解位于约束边界上，就可使用更有效的方案来找到最佳解。对于第 4 章中介绍的噪声消除问题，与观测图像相比，限制约束条件通常是平滑图像的卡方值。这通常假定是由经典噪声产生的，因此应有一个上限和一个已知分布。

对于层析重建，约束条件是满足射线积分方程。但它们并不完全一致，因此须对误差强加一个权重方案；线性加权是最简单和最常用的。业已证明，在多数情形下，约束条件允许的具有较高熵的解的聚类通常很难分辨。换言之，最大熵方法确实可以得到有用且鲁棒的解。尽管解仍然是迭代的，但与其他求解技术相比，这种方法仍是有效的。

13.6 重建图像中的缺陷

图 13.11 和图 13.15 中显示的重建例子，是使用通过计算仿真的投影数据计算的，它没有噪声或任何其他缺陷。在真实的断层扫描中，投影集中可能出现的各种缺陷会将误差传回到重建的图像中。使用相同的体模，可演示几个更常见的例子。

理想情形下，可使用大量的视角和沿每个投影集的足够检测器位置，来得到关于重建的足够信息。使用较少投影或视图的情形下，图像会具有更多的重建缺陷，很差的分辨率和边界，以及很差的精度和体素值的均匀性。图 13.19 显示了每组中较少投影的影响，但仍然使用了 180 个视角。重建后的图像以 100×100 个体素显示。理想情形下，这要求每个投影集中的射线积分数量至少等于宽度的 $\sqrt{2}$ 倍，或每个视图有 141 个方程。使用更少的视图时，重建的分辨率将会降低。

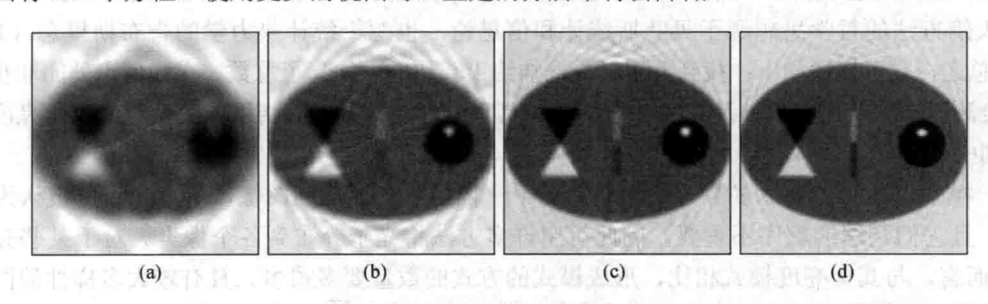


图 13.19　投影集中射线积分的数量对重建图像质量的影响。每幅图像都用 100×100 个体素重建，并由 180 个视角计算：(a)使用 25 个射线投影重建的图像；(b)使用 49 个射线投影重建的图像；(c)使用 75 个射线投影重建的图像；(d)使用 99 个射线投影重建的图像

如果使用的视角较少（但角度间距仍一致），那么重建中的缺陷会增加，如上面的图 13.15 所示。如果在整个范围内视角不是均匀间隔的，那么结果会更糟，如图 13.20 所示。

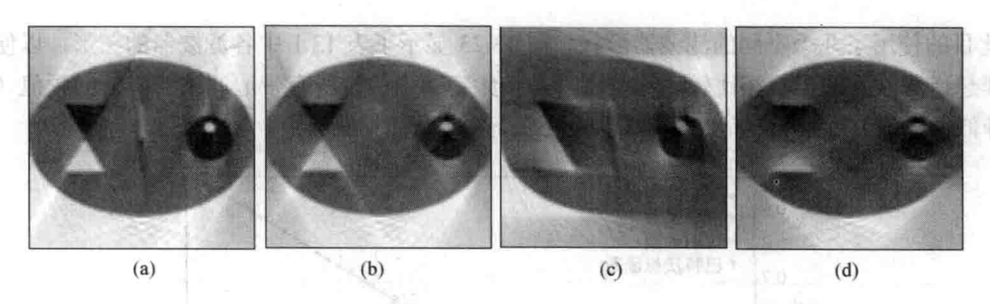

(a) (b) (c) (d)

图 13.20　使用未均匀覆盖角度范围的一组视角的效果：(a)150° 覆盖；(b)120° 覆盖；(c)90° 覆盖；(d)另一个 90° 覆盖

在真实的图像中，在投影集中的每个点处所检测到的 X 射线光子数，受到计数统计引起的波动影响。在许多情形下，包括医学和工业断层扫描，光子的数量是有限的。在医学应用中，对物体的总照射量进行限制是很重要的。在工业应用中，限制是 X 射线管或放射性同位素源的强度有限，并需要在合理的时间内尽可能多地获取视图。在任意一种情形下，所检测的 X 射线的数量变化呈近似高斯分布或正态分布，这种分布的标准差是统计得到的数字的平方根。统计 100 条 X 射线的平均值，可得到一个标准差为 10%($\sqrt{100}=10$)的变化，而统计 10000 条 X 射线的平均值，可将变化降低到 1%($\sqrt{10^4}=10^2$)。

重建过程会放大投影中噪声的影响。这种噪声影响与第 6 章通过反卷积消除模糊时的噪声影响相同，因此有着相同的数学成因。采用了使用整个 180° 数据的滤波反投影。该滤波过程抑制低频而保留高频，并统计波动的逐点随机变化数，进而显示在最高频率数据中。图 13.21 显示了结果。将百分之几的统计或计数波动添加到模拟的投影数据中，会在重建图像中产生更大的噪声。尽管体模的三个区域中，密度差的变化达 100%，但在投影数据中添加 10%或 20%的噪声后，有些区域完全消失了。

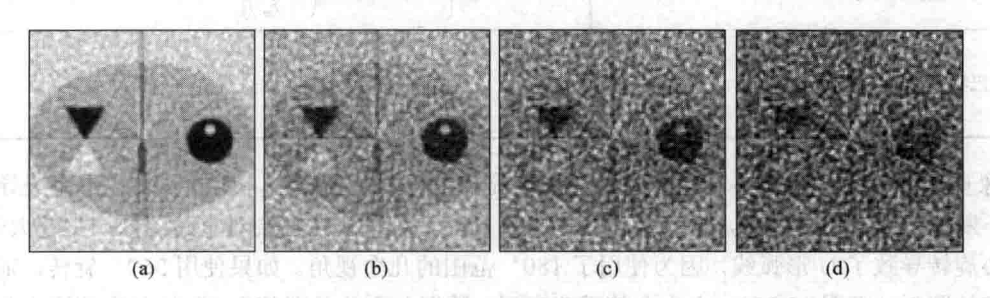

(a) (b) (c) (d)

图 13.21　计数统计对重建的影响。由已添加高斯随机波动的模拟投影数据重建的图像：(a)添加 2%；(b)添加 5%；(c)添加 10%；(d)添加 20%

通过滤波处理抑制投影数据的高频噪声，可在一定程度上降低噪声的影响，如图 13.22 所示。重建图像中的噪声变化降低了，但用于生成清晰边缘并揭示较小结构的高频数据也消失了。

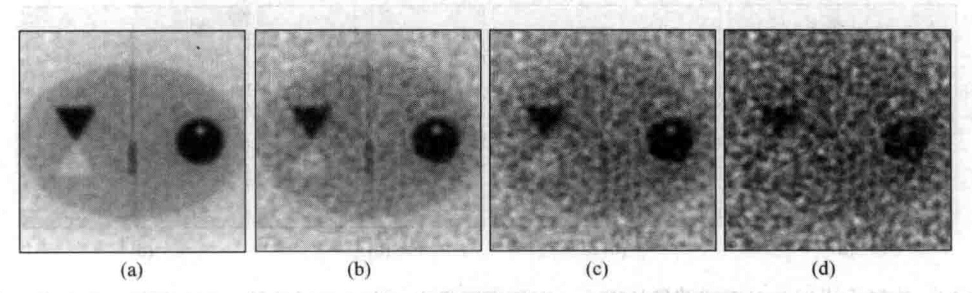

(a) (b) (c) (d)

图 13.22　采用图 13.21 中叠加的相同计数统计变化，但使用汉宁滤波器而非理想的逆滤波器来降低高频噪声，由相同的投影数据产生的重建

为此目的使用了几个不同的滤波器形状。图 13.23 显示了表 13.1 中各滤波器的图形，以便和上面描述的理想逆滤波器的形状相比较。这些图形是按频率画出的。所有带通滤波器降低低频值（防止模糊需要降低低频值），而所有降噪滤波器也会衰减高频，从而抑制噪声。

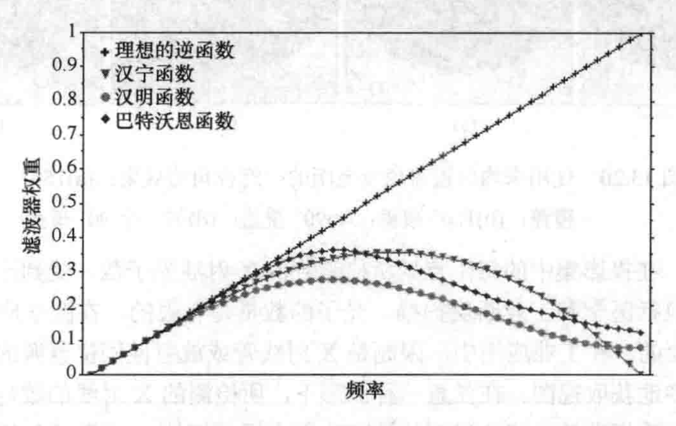


图 13.23　降低滤波反投影中噪声的（表 13.1 中所列）各个滤波器的剖面

表 13.1　图 13.23 所示反投影滤波器剖面

理想逆滤波器	权重 $=\lvert f\rvert$
汉宁滤波器	权重 $=\lvert f\rvert\cdot\left\{0.5+0.5\cdot\cos\left(\frac{\pi}{2}\cdot\frac{f}{f_m}\right)\right\}$
汉明滤波器	权重 $=\lvert f\rvert\cdot\left\{0.54+0.46\cdot\cos\left(\pi\cdot\frac{f}{f_m}\right)\right\}$
巴特沃恩滤波器（n 阶）	权重 $=\lvert f\rvert\cdot\frac{1}{1+\left(\frac{f}{2\cdot f_m}\right)^{2n}}$

图像重建中的另一个重要误差来源是不知道旋转中心的精确位置，或机械装置不完美导致了旋转中心不断变化（Barnes et al.，1990）。如图 13.24 所示，这种变化也会对重建图像产生较大的影响。偏离中心旋转导致了 U 形弧线，因为使用了 180° 范围的几个视角。如果使用 360° 旋转，则会出现一个完整的圆弧（见图 13.25），它也会使重建变形。数据并不总是在整个 360° 角度范围内收集，因为缺少偏离中心的旋转或射束硬化（在下面讨论）时，数据是冗余的。可变中心的影响和 360° 旋转一样大，但更难于识别。通常，旋转中心位置的恒定性比重建图像中期望的空间分辨率或体素尺寸小 1/10 左右。

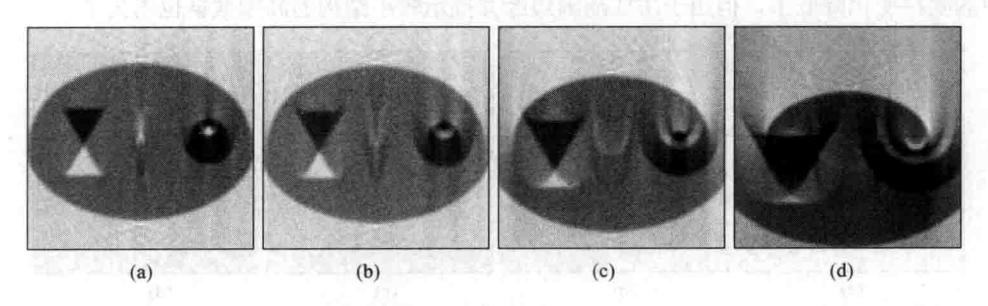


图 13.24　旋转中心误差对重建图像的影响。在每幅图像中，中心是一致的，但与旋转中假定的位置偏离了图像宽度的一个百分数：(a)0.5%；(b)1.0%；(c)2.5%；(d)5.0%

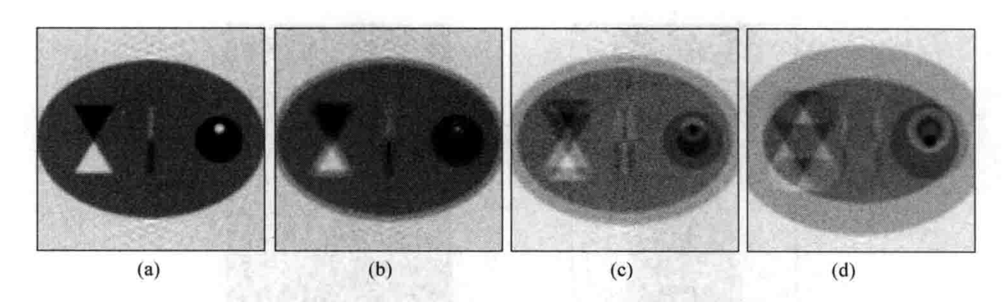


图 13.25　图 13.24 的再次重建，使用了相同数量的视图（180 个），但覆盖范围是 360° 而非 180°，旋转中心与假设位置偏离了图像宽度的一个百分数：(a)0.5%；(b)1.0%；(c)2.5%；(d)5.0%

类似地，在收集多个视图的过程中，要限制物体的运动。在医学成像中，最常见的成像几何是将患者放进一个圆形轨道内，发射源或检测器在该转道上旋转，这意味着扫描过程中患者必须摒住呼吸。如果期望图像显示心脏，则必须在比心跳更短的时间内一次完成所有视图的收集，或使用频闪触发方案在多次心跳的同一相对时间内收集图像。这两种方法都被用到。

13.7　射束硬化

射束硬化是指某个多色光源（如常规的 X 射线管）发射的低能量 X 射线被样本优先吸收的效应。其结果是体素的有效衰减系数是不同的，具体取决于该体素是在靠近源的一侧还是在远离源的一侧。沿该路径的这种变化如图 13.26 所示。射束硬化在医学断层成像中并不是什么大问题，因为人体各部分组成的变化很小。人体主要由水组成，有些部分含有一些碳元素或其他元素，如骨头中就含有钙元素。密度是变化的，重建图像中事实上也显示了这一点，但变化范围很小。这种均匀性就使得 X 射线管成为人们可以接受的源，并使得简单的滤波反投影成为一种合适的重建方法。

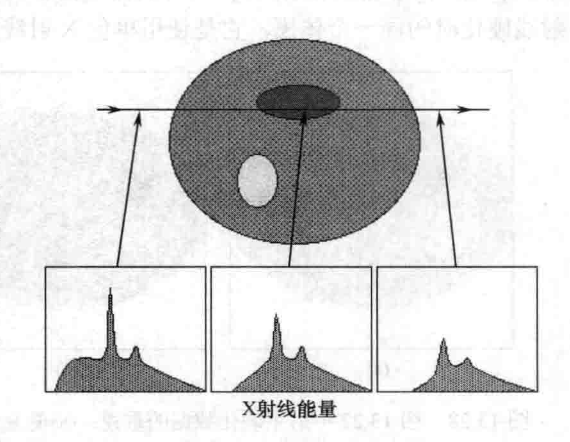
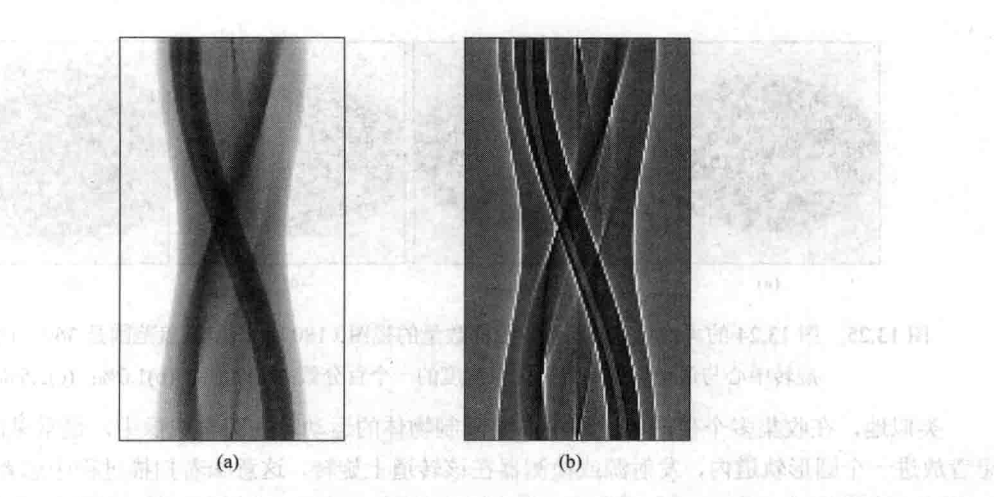
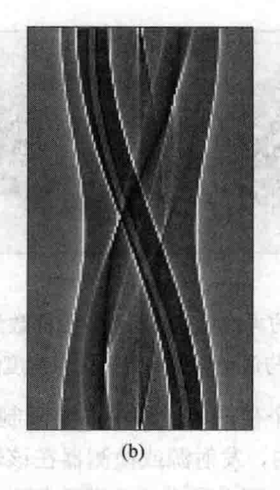


图 13.26　射束硬化示意图。在通过标本的路径的开始、中间和末端，显示了来自 X 射线管的 X 射线的能量谱。当较低的能量被吸收后，样本的衰减系数会发生变化，这种变化与任何实际的成分或密度变化无关

工业应用经常遇到成分变化很大的样本，这些成分包括整个元素周期表中的元素，以及从零（空隙）到十倍于生物组织密度的的物理密度。这种大范围的变化就使得射束硬化成为一个重要问题。一种解决方法是使用单色源（如放射性同位素或滤波后的 X 射线管）。另一种方法是使用两种不同的能量（Schneberk et al.，1991）或吸收与 X 射线散射数据的组合（Prettyman et al.，1991），并使用两个投影集来校正重建过程中成分的变化。但这种方法明显增加了复杂度，并需要一种代数求解方法而非反投影或傅里叶技术。

图 13.27 中显示了射束硬化对上面所用的同一个体模的影响。此时，样本成分指定为空（浅灰色的内部区域和周围区域）、钛（中灰色椭圆形大物体）和铁（小而暗的内部区域）。总宽度是 1 厘米，X 射线管的工作电压假定为 100 千伏。这是适度数量的射束硬化。更大的样本、更低的管电压、更高的原子序号元素、原子数或密度的更大变化，会产生更大的影响。

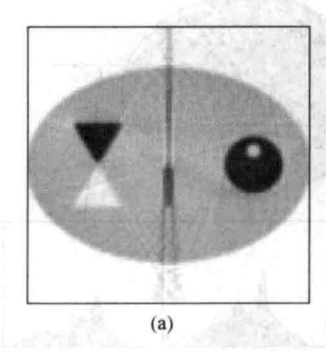




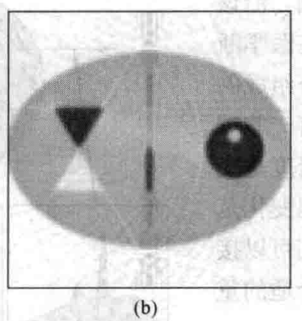
(a) (b)

图 13.27 (a)射束硬化对正弦图或 Radon 变换的影响；(b)射束硬化对逆滤波数据的影响。注意，每个特征的对比度会根据其在旋转物体内的位置变化

图 13.28 显示了分别使用涵盖 180° 和 360° 视角的体模的反投影重建。在多数断层摄影中，由于期望投影是相同的，而不考虑沿一条射线路径的方向，因此 180° 视角足以满足应用要求。在射束硬化情形下（如偏离中心旋转的情形），这种假设是不正确的，此时使用完整的 360° 数据会得到更好的结果。注意，伪影仍然存在。对中心的特征尤其如此，该特征中的较窄空隙难以看清。图 13.29 显示了无射线硬化时的同一个体模，它是使用单色 X 射线源产生的。



(a)



(b)

图 13.28 图 13.27 中射束硬化数据的重建：(a)覆盖 180° 的 180 个视图；(b)覆盖 360° 的 180 个视图

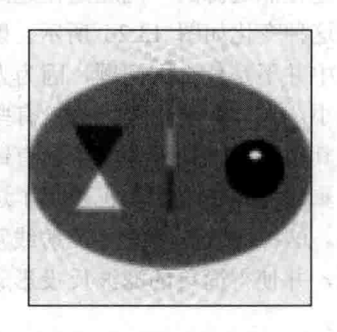


图 13.29 图 13.28 中体模的重建，此时使用的是单色 50 千伏 X 射线源。特别注意物体中心的空隙，它在图 13.28 中是不明显的

当 X 射线穿透材料时，降低透射强度的衰减系数由两个主要部分组成：激励束缚电子的 X 射线吸收，向某个不同方向一致或不一致的 X 射线散射。在任一情形下，光子都会在直射线路径上丢失，进而降低观测强度。但在散射情形下，X 射线可重定向到检测器阵列中的另一位置（见下面各代仪器设计的几何描述）。

发生这种散射时，测量的投影剖面中包含有额外的背景，背景上叠加了衰减数据。背景的出现也会在重建中产生伪影，如图 13.30 所示。这种影响看起来类似于射束硬化产生的影响。

此外，各个物体中的均匀区域是使用背景所致的一个变化密度重建的。图 13.31 显示了对一个简单的环形对象（骨骼截面的一个简化模型）的重建，图 13.32 显示了横跨重建中心的图形。重建中均匀密度的偏离称为杯突。这个例子使用了成分和密度与人体类似的材料。但人们并不总是要求通过医学断层成像来准确测量密度，而是仅显示内部结构和边界的位置。但在某些应用中并非如此：例如，射

线照相和计算机断层扫描就被用于测量骨质疏松症引起的骨密度损失。工业断层扫描经常要求准确测量密度以便量化处理（例如，压制粉末烧结的陶瓷零件）所致的部分梯度，因此人们很关注这种误差来源。图 13.33 显示了一个有趣的应用（Sirr & Waddle，1999），其中密度和尺寸的精确测量都很重要。

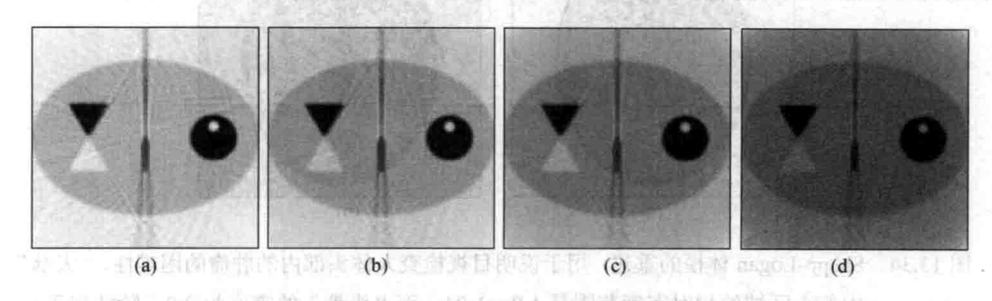


图 13.30　当测量的投影集包含平均亮度某个百分比的散射背景辐射时，图 13.29 中体模的重建：(a)5%；(b)10%；(c)20%；(d)40%

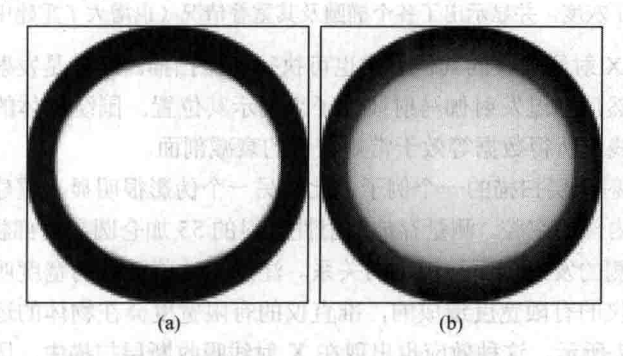


图 13.31　外部的碳酸钙（骨头成分）和内部的水（组织成分）构成的一个简单的环形的重建：(a)无散射背景；(b)投影数据中有 10%的散射背景

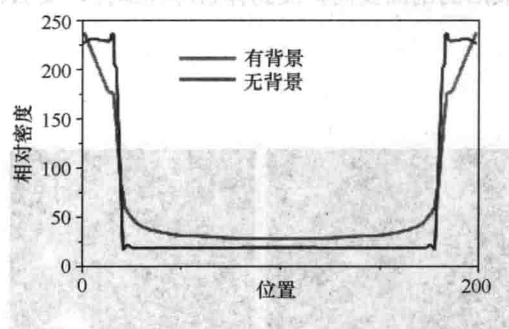


图 13.32　图 13.31 所示图像的密度剖面

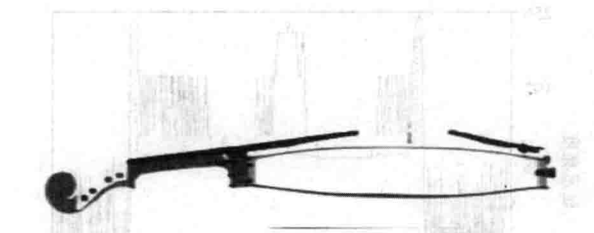
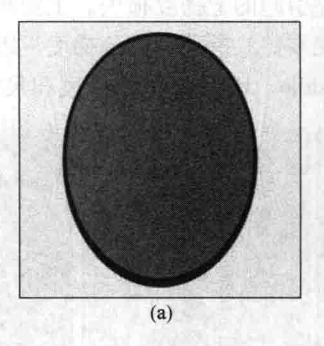


图 13.33　名为 Harrison 的小提琴的 CT 扫描。这把小提琴于 17 世纪制造于意大利克雷莫纳，它几乎完全是手工制作的，未经过任何“现代”加工

尽管医学应用几乎不需要精确测量密度，但却需要具有显示密度的较小变化的能力。常用于显示和评估此类性能的测试体模是 Shepp and Logan（1974）头部体模。该体模由密度接近于 1.0 的椭圆组成，是人体头部的简化形式，周围是致密的头骨，且包含有密码较低和较高的区域，以便模拟脑部结构和肿瘤。对这些区域成像的能力，对于在真实头部扫描中检测异常到关重要。

图 13.34 显示了这个体模的重建。使用全动态范围显示（值 0～255）来线性表示图像，并未揭示体模内的内部细节。应用直方图均衡（见第 5 章）扩展了直方图中心的对比度，以便能显示不同的区域。穿过结构中心的剖面图显示了具有均匀密度值的不同区域（见图 13.35）。


(a)

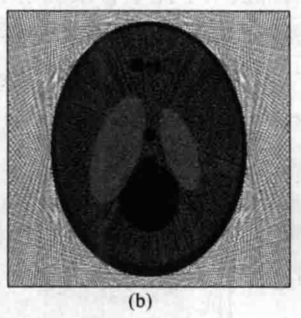

(b)

图 13.34　Shepp-Logan 体模的重建，用于说明目视检查人体头部内的肿瘤的困难性。“大脑”中各个区域的相对密度范围是 1.0～1.04，而“头骨”的密度为 2.0。除非应用一些对比度扩展，否则它们在重建图像中是不可见的(a)。直方图均衡(b)非线性地扩展了灰度，并显示出了各个椭圆及其重叠情况（也增大了重建中伪影的可见性）

如上所述，使用非 X 射线吸收的其他形式也可执行断层扫描。一种是发射断层扫描，此时放射性同位素放在物体内部，然后通过发射伽马射线光子来揭示其位置。围绕物体的并排检测器可以指定那些其上可放光子源的直线，所得数据等效于常规情形的衰减剖面。

图 13.36 显示了发射断层扫描的一个例子，此时另一个伪影很明显。重建中的明亮区域是含有放射性同位素的机构零件内部的空腔。测量存放放射性废料的 55 加仑圆桶内部物质的质量，使用了同一技术。正弦图显示了检测的发射剖面与视角的关系。注意，这些区域的宽度明显会随角度变化。这一变化是由检测器上准直仪的有限宽度造成的，准直仪的有限宽度会在物体的远侧生成一个能覆盖更宽尺寸的角度，如图 13.37 所示。这种效应也出现在 X 射线吸收断层扫描中，因为此时源和检测器上的孔径尺寸也有限。使用限制接收角的较窄准直仪也可降低测量的信号。如果已知准直仪的角度，那么这种影响可包含在重建中，要么逐步扩展数据，就如滤波后的剖面反向扩展到体素阵列那样，要么调整代数重建技术中的体素权重。

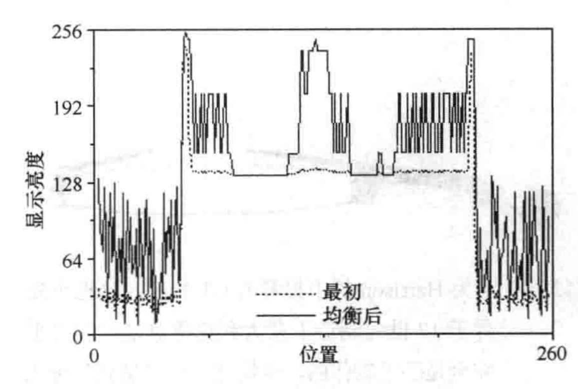


图 13.35　横跨图 13.34 中图像的亮度剖面（显示了各区域的一致性和清晰性）和直方图均衡的影响（增大了噪声的可见性）

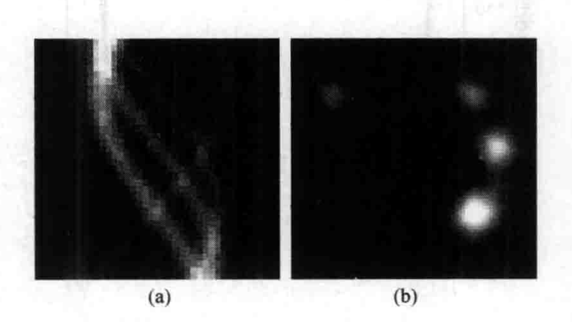

(a)　(b)

图 13.36　发射断层成像。样品是铝块，其内部存在几个含有放射性钴的圆柱空腔。针对每次旋转，检测器收集一系列亮度剖面，并显示为正弦图(a)。注意，由于入口准直仪到检测器的角度大小有限，因此每个圆柱体轨迹的宽度会随样品的旋转而变化。截面的重建显示在(b)中

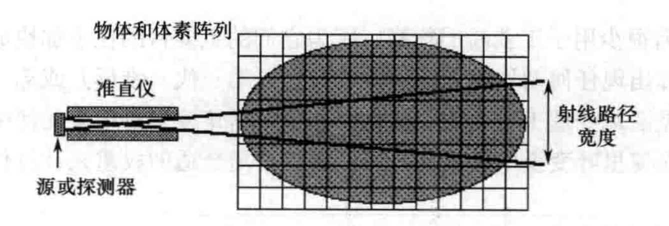


图 13.37　准直仪的有限角度对物体不同部分的尺寸和体素的影响

13.8　成像几何

第一代断层成像系统通过移动发射源和检测器来在不同的视角收集投射集，如上面的图 13.8 所示。收集投影数据的完整过程如图 13.38 所示。这就是所谓的锥形射束或平行波束几何，其中每条射线积分是平行的，且投射集可直接反投影。由于只能以较小的立体角来发射 X 射线，而且此时只能使用一个检测器，因此效率不高，但仍然用在一些工业成像情形中，因为此时人们关心的并不是获取数据的时间。

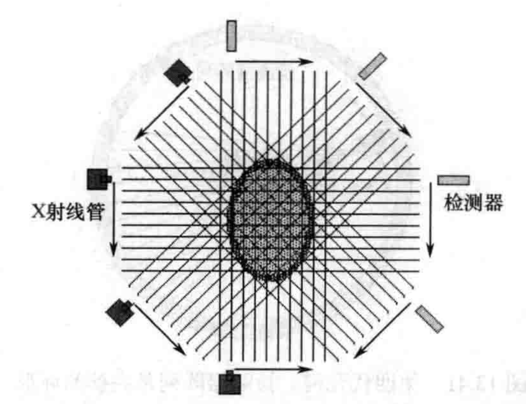


图 13.38　第一代几何。在每个点处同时移动检测器和信号源来收集每个投影集，并旋转到许多视角来收集所需的全部数据

第二代设备增加了一组检测器，以便可检测扇形 X 射线束，且同时可沿几条直线测量衰减，如图 13.39 所示。这一过程收集相同数量的数据时，所需的视角较少，但对于不同的角度和某些数据排列，需要来自每个检测器的衰减测量数据。

扇束几何的效率较高，而第三代成像设备中所用的下一个逻辑步骤，则使用了较大的检测器阵列（这些检测器按弧形方式排列，以便每个检测器都能覆盖 X 射线法向入射的相同角度）。激励 X 射线管时，检测器和射线管围绕物体旋转，以产生一系列的视角（见图 13.40）。在第四代系统中，则安装了环形检测器，且只有发射源旋转（见图 13.41）。此时，X 射线不再垂直入射到检测器中。第五代系统中几乎不移动硬件，它通过使电子束针对固定的目标环发生磁性偏转来产生 X 射线，而通过旋转 X 射线源来产生如同第四代系统中的相同几何，其曝光时间更短，且无须移动机械零件。

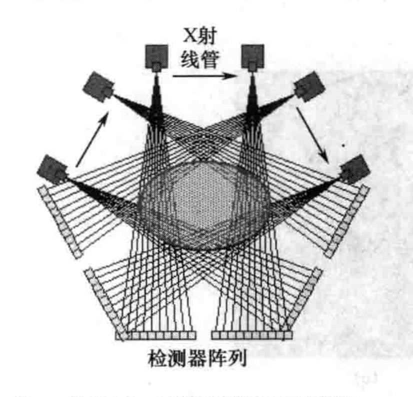


图 13.39　第二代几何。检测器阵列同时测量扇形束中的衰减，与第一代系统相比，收集数据所需的视角更少

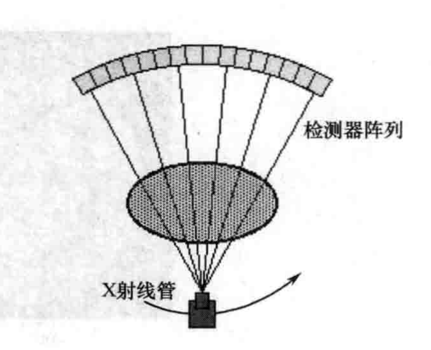


图 13.40　第三代几何。X 射线管和检测器阵列围绕待成像物体一起旋转，同时快速激励射线管产生每个视角

后面几种成像几何很少用于工业断层扫描，因为它们的主要目的在于加快成像速度、缩短曝光时间，进而在被成像人体出现任何动作前获取所有的投影。第一代（锥形）或第二代（扇束）方法收集一系列离散视图，可灵活地处理工具问题。但是，如果根据角度来对扇束几何中各个检测器的射线积分排序，此时不管是在傅里叶变换中使用反投影，还是使用合适的权重来计算代数重建，这些方法都是等效的。

有些成像技术会使用不同的几何和重建。例如，我们熟悉的超声波图像（主要用于检查子宫内胎儿发育情况或检查血液流动来监控心脏问题）是由扇束剖面得到的（见图 13.42），其所接收的信号是由内部表面反射回来的信号。处理这些剖面即可得到这些表面的视图，详见下一章中的介绍。正电子发射断层扫描（PET）、单光子发射光谱仪（SPECT）和磁共振成像（MRI）的几何不同于 CT 扫描，它们记录的是内部结构的不同信息。

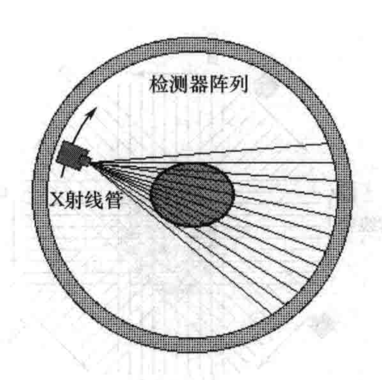
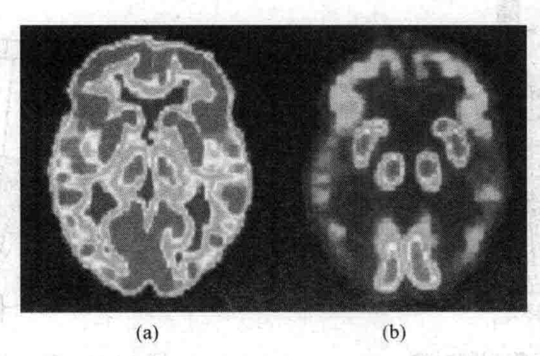


图 13.41　第四代几何。检测器阵列是完整的环形且固定。X 射线管围绕物体旋转并激励。顺序处理检测器的输出就可得到投射集



图 13.42　11 周大胎儿的超声波图像

PET 和 SPECT 由某种原子同位素的衰减来局部化该同位素。在 PET 中，原子核产生的正电子有一组 511 千电子伏特的光子，这两个光子的运动方向完全相反。使用检测器环检测它们，可在两点间的直线上的某个位置找到原始同位素标记。PET 扫描最常用于检测癌症，并通过描述肿瘤中的生物化学变化来检查癌症治疗的效果。人们还使用 PET 来了解大脑活动，进而研究代谢下降与老人痴呆症的关系（见图 13.43），以及检查心脏病发作后的心肌损坏情况。PET 扫描中所用的放射性元素通常是具有较短半衰期的同位素，如碳 11、氮 13、氧 15 和氟 18（半衰期分别为 20 分钟、10 分钟、2 分钟和 110 分钟）。



(a)　　(b)

图 13.43　显示了代谢下降与晚期老年痴呆症间关系的 PET 扫描图像：(a)正常的成人大脑；(b)晚期老年痴呆症患者的大脑

SPECT 使用准直检测器检测来自同位素衰减的伽马射线。最常用的同位素是锝 99（半衰期为 6 小时）和碘 123（半衰期为 13 小时）。检测器阵列从一组平行的投影收集信息，并围绕物体进行扫描来得到多个视角的信息。SPECT 的主要用途是了解到达心肌的血流情况（见图 13.44），进而检测机能不全和冠状动脉疾病区，它是运动前后获取图像进行压力测试的一部分。与 MRI 或 CT 方法相比，PET 和 SPECT 的分辨率都较差，但仍足以用于诊断和研究目的，尤其是与其他高分辨率的成像形式结合使用时。PET 和 MRI 数据与 X 射线断层扫描得到的密度信息高度互补，且在许多情形下，通过在配准中叠加这些剖面图像，可组合由不同方式得到的扫描图像（见图 13.45）。对具有不同分辨率和体素大小的这些图像进行配准，并显示不同结构的对比度也并不简单（Hajnal et al.，2001）。

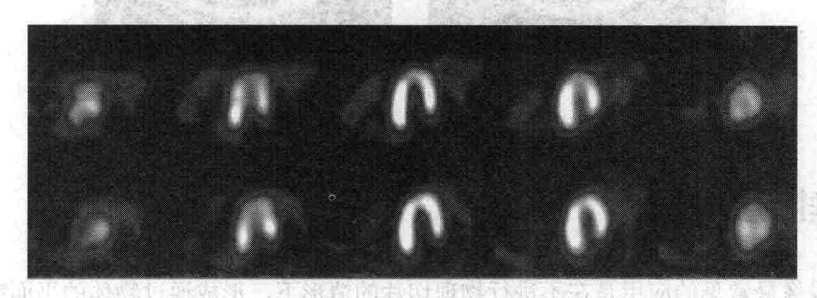


图 13.44　显示心脏血液流动的 SPECT 图像。这些图像是平行截面，颜色和亮度表示同位素的浓度

质子放到磁场中后，它们会以与磁场强度相关的某个频率振荡（共振），并以该振荡频率吸收能量。当质子回到基态时，能量就会被再次辐射出去。松开与磁场平行且垂直的磁化组件（它们具有不同的时间常数 T_1 和 T_2），就会发生再次辐射。MRI 的信号强度取决于质子浓度（对于医学成像，是组织中的水浓度），但对比度取决于 T_1 和 T_2（如图 13.46 所示），质子浓度和对比度都受流体黏度或组织硬度的强烈影响。加权组合这两个信号，即可控制所观察的图像。

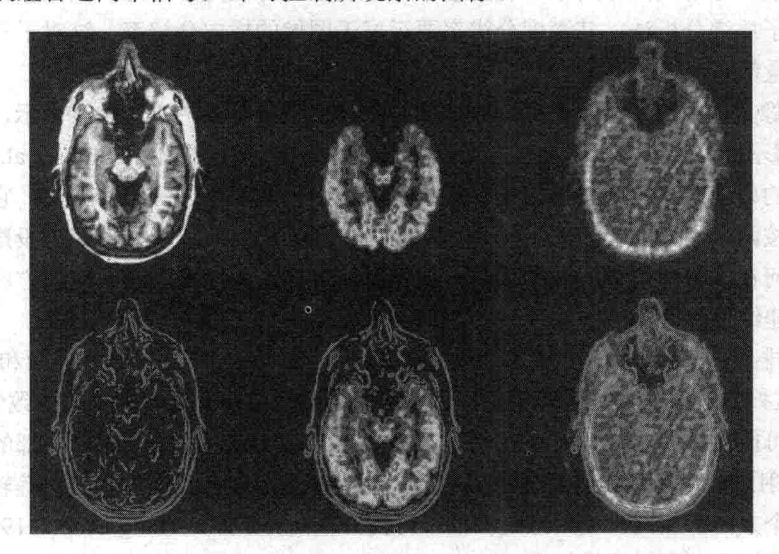


图 13.45　PET 和 MRI 扫描图像的配准

MRI 图像重建时，需要精确校准跨越被成像区域的磁场梯度，以便局部化不同的频率。射频信号的傅里叶变换可提取质子的位置，在许多方向收集信号及将它们重组为二维或三维空间分布的原理类似于 CT。MRI 可对人体中的所有软组织进行可视化成像，而在 CT 中软组织通常不能产生较好的对比度（尽管在特定情形下可让患者服食造影剂或对患者注射造影剂）。报道称，结合 CT 与 MRI 图像可为某些医疗诊断提供帮助（Bartling et al.，2005）。

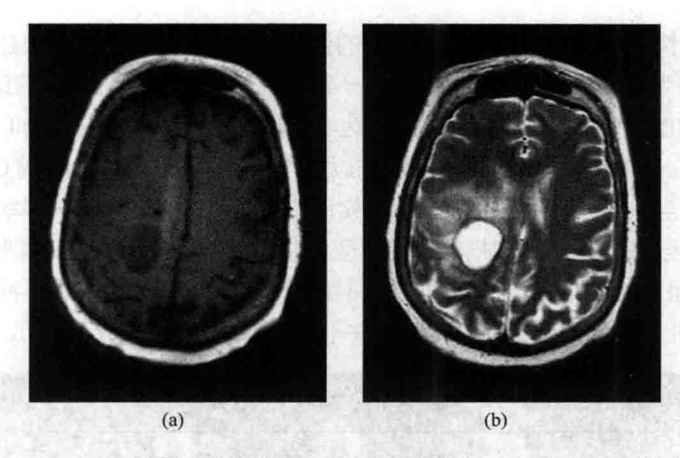


图 13.46 (a)使用 T_1 加权的 MRI 切片；(b)使用 T_2 加权的 MRI 切片

13.9 三维断层成像

尽管断层成像最常见的应用是在不进行物理切片的情形下，形成通过物体的平面截面，但该方法可直接生成完整的三维图像。第 14 章展示了体数据三维显示的几个例子。多数此类图像，包括许多断层成像图像，实际上都是连续切片图像。不论是由物理切片、光学切片（如使用共焦光学显微镜）形成，还是切片的传统断层成像重建，这些方法并不是理想的三维数据集。

区别在于，每个图像平面中的像素是方形的，但随着它们在三维情形下扩展为体素，就不一定会是立方体了。平面间的距离或深度分辨率，并不等同于像素大小或平面内的分辨率。事实上，这些方法的深度分辨率很少等于横向分辨率。有些技术，如物理切片或光学切片，其深度分辨率很低。其他技术，如二次离子质谱分析法，其深度分辨率要远好于图像的横向分辨率。这对于三维图像表示、图像处理和三维测量都有深远的影响。

采用层析成像重建，可实现真正的三维成像。物体由三维阵列的立方体素表示，各个投影集则成为二维阵列（投影图像）。每个投影来自一个点，且被称为锥形射束几何（Shih et al.，2001），类似于单切片投影所用的扇束方法。观察方向集中必须包含移出平面和移进三维的方向，它们由两个极角指定。这并不要求按两个不同的极角来旋转物体，因为使用锥形射束成像几何可为投影直线提供不同角度，就像扇束几何在二维情形下所做的那样。但使用尽可能均匀地覆盖三维空间方向的一系列视角，可得到最好的重建结果。

有几种几何排列方式。最简单的几何排列是，关于某个轴旋转样品（或发射源和检测器阵列），如图 13.47 所示。这种方法的优点是旋转精确，因为重建的质量取决于旋转中心的一致性。另一方面，重建体素中的伪影可能会很明显，尤其是在平行于轴的方向以及靠近样品顶部和底部的地方。单轴旋转方法最常用于 X 射线、中子或伽马射线断层扫描，因为样本可能会相当大且相对等轴，以便辐射穿过样品的距离在每个方向都相同。使用螺旋扫描，可提升轴向分辨率（Wang et al.，1991），在这种扫描方式中，轴向移动标本的同时旋转标本（见图 13.48）。对于小到中等尺寸的物体，这是首选的几何排列方式。

对于电子断层扫描，多数样品是薄切片，且几乎没有透射电子显微镜台允许样本关于其轴整体移动。对于板状样本，采用的几何排列方式是锥状投影的一系列倾角（Barth et al.，1989），如图 13.49 所示。控制样品在容器中的倾斜和旋转，进而以足够的精度收集这些投影来进行良好的重建较为困难。有些 TEM 样品由许多重复的结构（大分子、病毒颗粒等）组成，因此样本的单幅图像就可收集足够的

不同投影视图来进行重建。由于使用了许多具有不同方向的各个（但大概相同的）对象，与使用等角增量相比，这种方法称为随机投影。TEM 中非常小的光束孔径角，可产生基本平行的投影，而非锥形射束投影，这就简化了重建并使得反投影很简单。但在圆锥中使用有限的视角会产生伪影，因为轴向的有用信息很少（有时也称为丧失锥信息）。Frank（1992）和 Kubel et al.（2005）深入回顾了电子显微镜层析成像的当前进展。

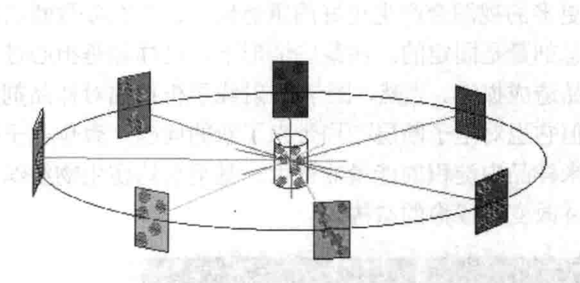


图 13.47　单轴旋转样品获得径向锥形射束投影，进而由这些投影进行体积成像的几何排列

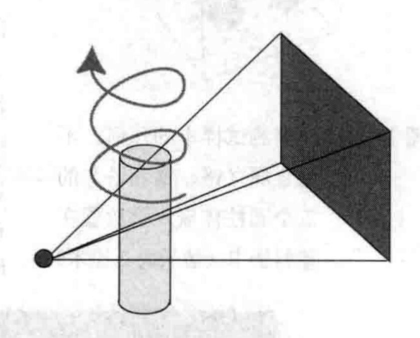


图 13.48　在螺旋扫描中，样本旋转的同时被抬升，因此多个锥形射束投影就测量了沿不同角度通过体素的吸收

从理论角度看，对于给定数量的投影，当它们均匀分布在三维空间中时，可得到最好的重建结果（见图 13.50）。很难构造一种机械装置来关于两个中心轴精确地旋转，且很少使用这种技术。



图 13.49　在圆锥形排列中倾斜样本可产生一系列平行束投影，进而用于透射电子显微镜重建。沿圆锥轨迹的间距不必一致

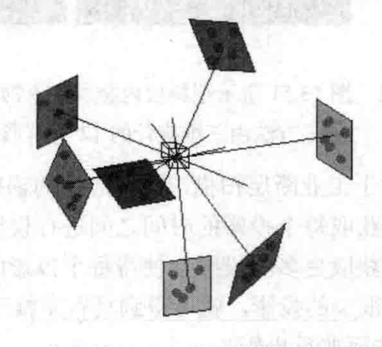


图 13.50　使用关于两个轴旋转样本得到的一系列三维投影，可得到最优的三维重建

三维重建可以使用任何二维重建方法来完成。对于傅里叶反变换，频率空间同样是一个三维阵列，每个投影产生的二维图像被变换，且复数值被画在阵列中的各个平面上。对于二维情形，需要尽可能完整且均匀地填充空间。傅里叶变换是在三维空间进行的，但它是低维空间中方法的直接扩展，且该变换可在一维情形下一次性完成（相继沿 u, v 和 w 方向）。

对于三维重建也可使用反投影，如在二维情形中那样，反投影是傅里叶变换的一种数学实现。二维图像的滤波必须由二维卷积进行，而二维卷积要么在空间域中由核运算实现，要么在傅里叶域中由乘法实现。反投影方法的主要缺点是，对于锥形射束几何而言，权重矩阵的计算很复杂，尤其是在结合使用螺旋扫描时。这些值代表了每个射线积分经过每个体素的衰减路径长度。反投影需要大量的视图才能避免伪影，它常与单轴旋转扫描或螺旋扫描结合使用（Feldkamp et al.，1984；Smith，1990；Shih et al.，2001）。应用完整的三维锥形射束集合是困难的，因为它们以相对较大的角度分隔。

代数重建（ART）方法也适用于体素阵列。获得一组均匀视角的困难性，会使得 ART 方法比傅里叶

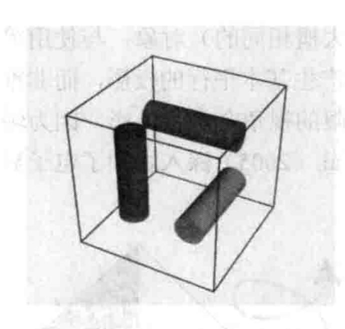


图 13.51　一个测试样本的几何。不同金属（铬、锰和铁）的三个圆柱体嵌入物放置在塑料块中（边长为 2 厘米）

反变换方法更有吸引力。事实上，在使用诸如 ART 的一种迭代技术时，较少数量的视图就可获得较好的效果。图 13.51 和图 13.52 显示了一个例子。标本（边长为 2 厘米）是一块塑料，塑料中有三个不同的金属圆柱体，这些金属分别是铬、锰、铁，它们的密度基本相同。使用低功率工业 X 射线源，由三维旋转的 12 个视图得到的层析成像重建，清晰地显示出了这些嵌入物（Ham，1993）。

当然，更多的视图会产生更好的重建结果。但在多数断层扫描情形下，总剂量是固定的。在某些情形下，这样做是担心过量辐射会对样品造成损害。当然，医学 X 射线层析扫描对样品剂量也有限制。但它也对电子断层扫描带来了新的问题。聚焦电子束在每立方纳米样品中淀积的能量非常大，甚至会灼伤生物组织，破坏分子，并改变待成像的结构。

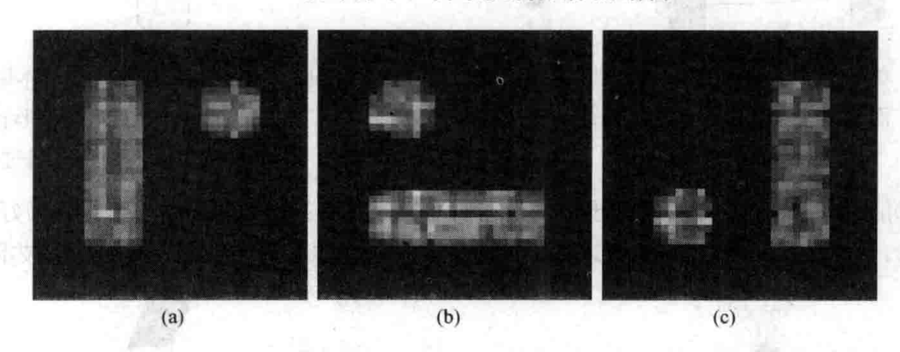


图 13.52　图 13.51 所示塑料块内金属嵌入物的 *xy*、*yz* 和 *zx* 方向的三个平面的层析成像重建，它是使用 ART 方法由三维旋转的 12 个锥形射束投影重建的

即使对于工业断层扫描，可以产生的辐射总量获取图像所需的时间也是有限的。因此，就需要在投影数量和获取每个投影的时间之间进行权衡。在每个投影上花费更多的时间可改善视图图像的统计质量，因此获取更多的投射会使得每个投影的噪声更多，反之亦然。在具有有限总光子预算的某些情形下，使用很少的投影，就可得到具有完整三维旋转的最佳质量重建（Ham，1993）。这种方法需要一种 ART 方法而非反投影。

使用低强度源时，有限数量的光子会变得非常关键。同步加速器（Helfen et al.，2003；Betz et al.，2007）是一种非常好的 X 射线源，其亮度高，具有选取某种特定单色能量的能力，但通常并不用于断层成像。伽马射线放射源方便携到，但处理起来很困难，且强度较低。X 射线管是断层成像的一种方便放射源，其电压可调，并提供许多种可发出不同 X 射线光谱的靶材料。射线管不是单色的，因此对许多标本而言，这可能会导致明显的射束硬化。

也可使用吸收滤波器来从多色源中选取单波段能量。对于每个视角，使用滤波器收集的两幅投影图像，其吸收边缘能量是不同的。两幅图像之比可产生其吸收边缘位于两个滤波器能量之间的元素的衰减信息，如图 13.53 所示。一系列这样的图像对可提供许多元素空间分布的单独信息。图 13.54 给

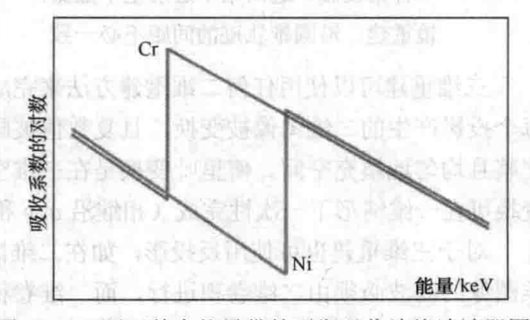
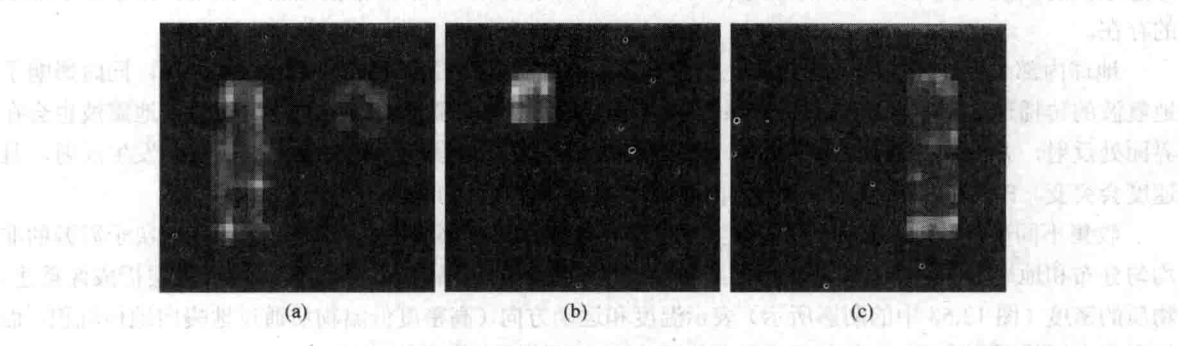


图 13.53　隔开单个能量带的平衡吸收边缘滤波器图。该图表明，作为两个不同滤波器的能量函数的吸收系数，包含了铬和镍两种元素。样本中吸收边缘介于这两个能量之间的元素（锰、铁和钴），按两个强度之比来成像

出了一个例子，其中使用滤波器选取了图 13.51 所示样品内三个金属嵌入物中的两个。使用滤波器降低了 X 射线管的已经很低的强度，且使用两幅图像之比给出了投影统计质量的进一步限制。因此，使用少量视图得到最可能的投影图像就很重要。这该例中，用到了绕轴三维旋转的 12 个投影。图 13.55 表明单轴旋转并记录相同数量的视图时，重建结果中出现了明显的伪影。



(a) (b) (c)

图 13.54 图 13.52 所示三个相同平面的重建，但它使用了通过不同滤波器（铬箔和铁箔）的两个投影之差得到的图像，这就形成了一个带通滤波器来选取一个较窄频段的 X 射线能量。注意，嵌入物（铬）被排除在外，但锰和铁嵌入物是可见的。重建使用了三维旋转的 12 个圆锥射束投影

电子显微镜所产生的图像，其对比度是由衰减造成的，由不同角度的一系列视图，可重建出显示三维结构的图像。由于晶体结构原子平面的电子衍射，因此使用任意系列角度难以实现材料标本的重建。计算普通的衰减，很难模拟这种放射源的对比度，因为体素在不同的方向会有不同的值。但对于非晶体材料（如生物标本），重建则很简单（Engel & Massalski，1984；Hegerl，1989）。

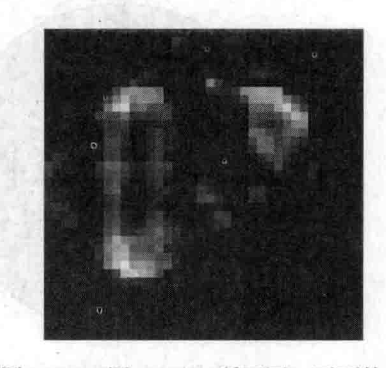
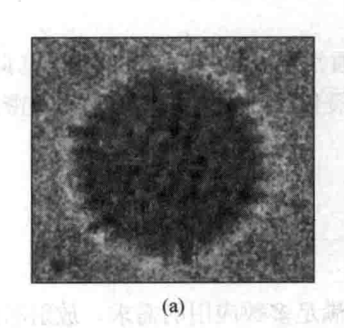
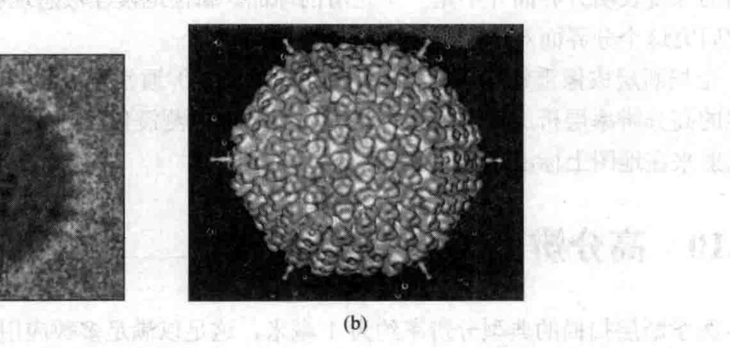


图 13.55 图 13.54(a)所示同一平面的重建，重建方式与使用两个滤波器的方式相同，但此时使了 12 个共面投影（仅关于一个轴旋转样品）。注意嵌入物之间和嵌入物内部增强的伪影

比使用单个标本的多个方向来收集一系列不同视图更有效的是，使用碰巧具有不同方向的同一标本的许多不同图像，如上所述。图 13.56 显示了一个例子。二维图像是单个病毒颗粒的电子显微照片。标本是引起呼吸系统疾病的一种腺体病毒。

防止标本损坏的低剂量电子使得该图像中有很多噪声。但在典型的标本中，有许多这样的颗粒，且每个颗粒的方向都是不同且随机的。通过收集各幅图像，并参考病毒表面上各个三角面的位置来对每幅图像的方向进行索引，可三维重建颗粒，其中每个体素值是电子宽度。对病毒的蛋白质外壳表面建模，可生成图 13.56(b)所示的表面渲染图像（Stewart & Burnett，1991）。





(a) (b)

图 13.56 (a)单腺体病毒颗粒的透射电子显微镜图像；(b)由许多随机方向透射图像中的腺体病毒颗粒表面实现的三维重建。注意嵌入物之间和嵌入物内部增强的伪影

以完全不同的尺寸，人们使用地震测定法对地球本身进行了断层成像。地震波由地震或大爆炸（如核武器试验）产生。这样的大事件产生两种类型的波，这些波通过地球内部后，由地表上许多不同位置上的检波器接收。P 波（压力波）是可穿透地球内部任何部分的纵波，S 波（剪切波）则是无法穿透液态地核的横波。事实上，英国地震学家 R. D. Oldham 在 1906 年就已根据地震 S 波形图推断出了液核的存在。

地球内部温度、压力和成分的变化，导致地震波的传播路径是弯曲的（见图 13.57），同时影响了地震波的传播速度，就像玻璃的折射系数导致光在镜头中弯曲那样。与光的性质类似，地震波也会在界面处反射，并导致传播速度的突变。地壳和地幔、地幔和地核的分界面处，地震波会发生反射，且速度会突变。P 波和 S 波的传播速度不同，因此对成分和密度的响应也不同。

收集不同事件的地震记录，可创建并不均匀覆盖地球的一组射线路径，这组路径取决于震源的非均匀分布和地震仪的分布。但对沿不同路径穿透地球的波的传播时间进行分析，可形成层析成像重建。物质的密度（图 13.58 中的阴影所示）表示温度和运动方向（高密度低温物质通过地幔向地核沉积，而低密度高温物质则上涌）。火山活动和大陆漂移导致了地幔中物质的流动。

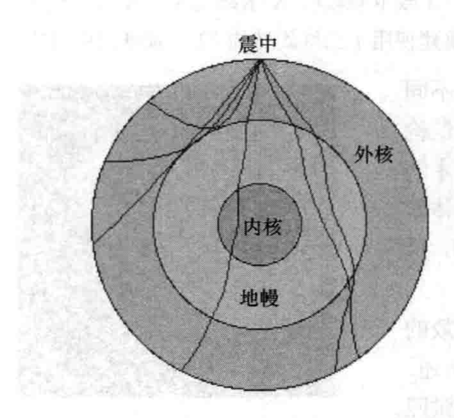


图 13.57　地震产生的压力波和剪切波的传播路径图（未按比例画出），该图揭示了经由地核和地幔的传播路径上的密度信息与不连续位置

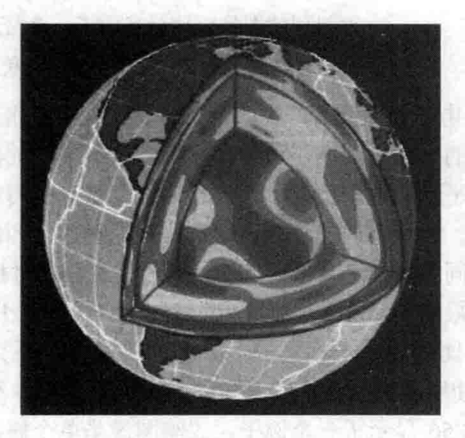


图 13.58　计算出的地幔断层扫描图像，显示了岩石密度的变化（明亮阴影处的温度高，密度小的岩石位于上方，而密度大的岩石则位于下方）

有用的波是从各分界面反射回来的波（一次反射波或多次反射波）。例如，对于直接到达检波器的 S 波与从地核与地幔分界面反射回来的 S 波，由它们的传播时差可求出分界面的深度，误差小于 1 千米，求出的深度表明分界面并不是一个光滑的球面。黏性地幔在液态地核上的流动，就产生了地球的磁场，因此研究这个分界面对于了解地球动力学很重要。

全局断层成像重建对结构的较小细节（如断层）通常不敏感。正在加州地下（此处存在许多断层）开展的高分辨率层析成像项目，采用了高灵敏度的测震仪器阵列，并使用了非常密集的小地震所产生的反射来在地图上标出断层的位置。

13.10　高分辨率断层成像

医学断层扫描的典型分辨率约为 1 毫米，这足以满足多数应用的需求。放射科医生通常使用具有标准化方向的平面切片图像，来识别正常特征和异常特征。但人们也会使用真实的三维层析成像来研究各种材料（包括金属、陶瓷、复合材料、聚合物和较大的工业零件）的微观结构。有些结构特征不

能由传统的二维显微截面确定，包括确定三维体中不同形状颗粒的数量、网络拓扑或孔隙结构，而这些结构决定了流体对这些材料的渗透率。

这种信息只能通过三维数据集来确定，三维数据集须具有足够的分辨率和理想的立方体素。使用同步加速器作为非常明亮的X射线点源时，分辨率可优于1微米。使用更稳定的放射源（如微焦X射线管）时，也可得到类似的分辨率。对这样的放射源进行滤波，可对特定的元素成像，如上所示。

锥形射束几何适合于这种类型的显微结构成像，因为它可放大结构（Johnson et al.，1986；Russ，1988；Kinney et al.，1989，1990；Deckman，1989）。图 13.59 示意性地显示了这一特性。放大是严格几何意义上的，因为 X 射线不会被镜头折射，但放大率可高达 100:1。射线经磷光体/通道板和合适的强化转换为可见光后，使用常规的固态检测器就可以收集投影的图像。由于常用的小光斑 X 射线源的强度很低，因此需要使用高亮度放射源（比如在一个同步加速器中可供使用的）来进行高分辨率成像。对于天文成像，则需要对使用相同的冷却式电荷耦合元件（CCD）摄像机拍摄的图像进行平均操作。

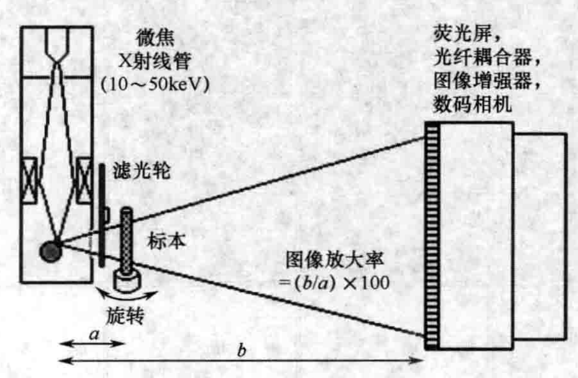


图 13.59 锥形射束成像系统图解。投影图像放大率是比值（b/a）。可以达到的分辨率受限于 X 射线源的光斑大小和检测器的空间分辨率

三维成像需要很多体素，而重建过程则需要大量计算。重建所需的时间通常仍然要短于收集各幅投影图像的时间。如上所述，生成这些图像的光子通常是有限的，且大量噪声会对重建造成影响。在视角有限的情形下，为了从强度有限的放射源收集质量较好的投影，要求在三维空间中合理地安排视图，以最优地覆盖这些极角。这种安排要求旋转和倾斜机构具有较高的质量，因为旋转中心必须恒定并且使变动保持在几微米范围内，进而保证图像的质量。螺旋扫描通常更容易实现此目的，因此人们更常使用这种扫描方法。

展示三维信息需要广泛使用计算机图形学方法，如第 14 章所示。图 13.60 显示了几个体素的一组平面，这些体素是多孔氧化铝陶瓷体的三维断层重建。各个颗粒是直径约为 100 微米的球体，它们约占样品体积的 60%。各个体素是大小为 10 微米的立方体。图 13.61 显示了通过样本的一个投影视图，它是一幅二维图像，图像中的球形颗粒沿视线重叠，且是部分透明的。这一数据的三维展示如图 13.62 所示。

自从本书第一版（1990 年）中给出了来自一个实验装置的这些例子后，技术一直在进步，并出现了几种这样的商用设备（Wang & Vannier，2001；Chappard et al.，2005）。图 13.63 显示了木材的一幅图像，该图像的分辨率可媲美光学显微镜图像的分辨率。这项技术通过创建亚微米分辨率的三维显微图像，已广泛用于多种应用中，包括微电子器件检查、材料检查及其他检查。1000×1000 大小体素的 1000 个切片的重建，1990 年时需要一台 Cray 超级计算机，而现在的台式计算机就可完成这一任务。图 13.64 显示了所遇到的样品和展示类型的其他几个例子。

图 13.60　重建体素的 20 个平面，显示了由直径为 100 微米的球体组成的烧结氧化铝陶瓷

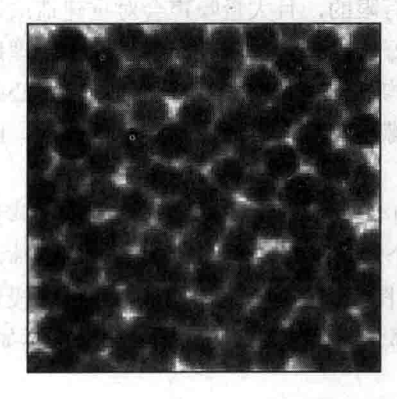


图 13.61　通过图 13.60 所示结构的一个二维锥形射束投影集

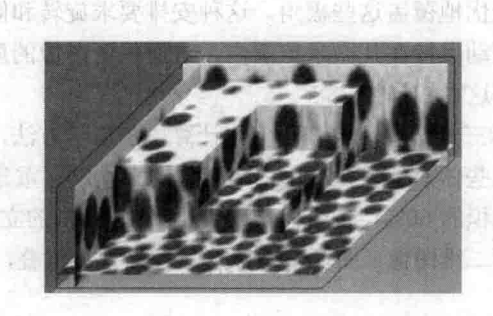


图 13.62　图 13.60 中数据的三维展示，垂直方向已拉伸

各种层析成像方法的绝大多数应用都会产生可视化图像［要么是一组二维切片，要么是三维渲染（如第 14 章所示）］，进而进行人工检查和解释。此时会执行一些定量测量（Hanke，2003；Rangayyan，2005；Leonard et al.，2007），但使用第 9 章中介绍的体视学方法来获得表征三维结构度量性质的数值更为有效。

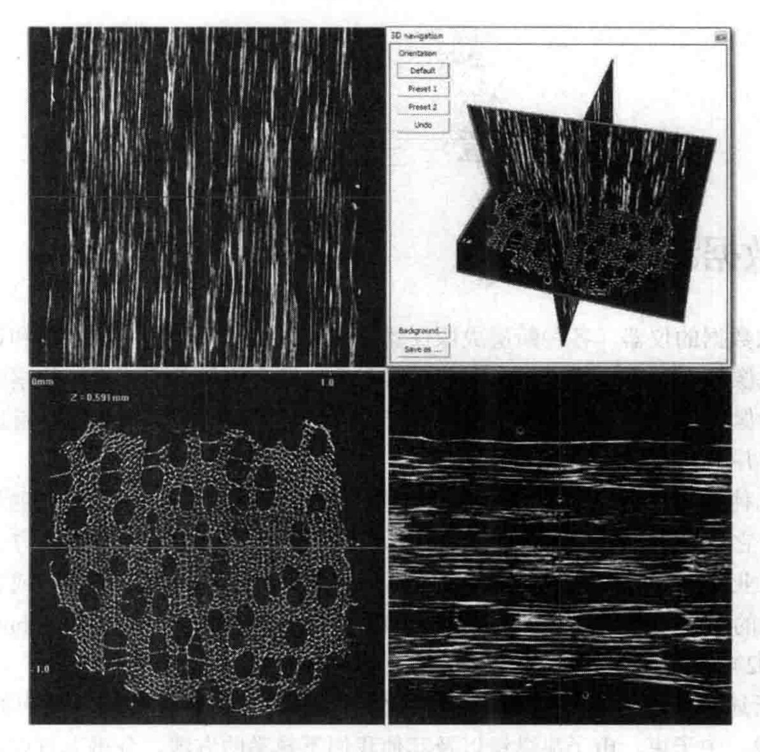


图 13.63　显微层析图像，显示了通过木质茎的三个正交切片

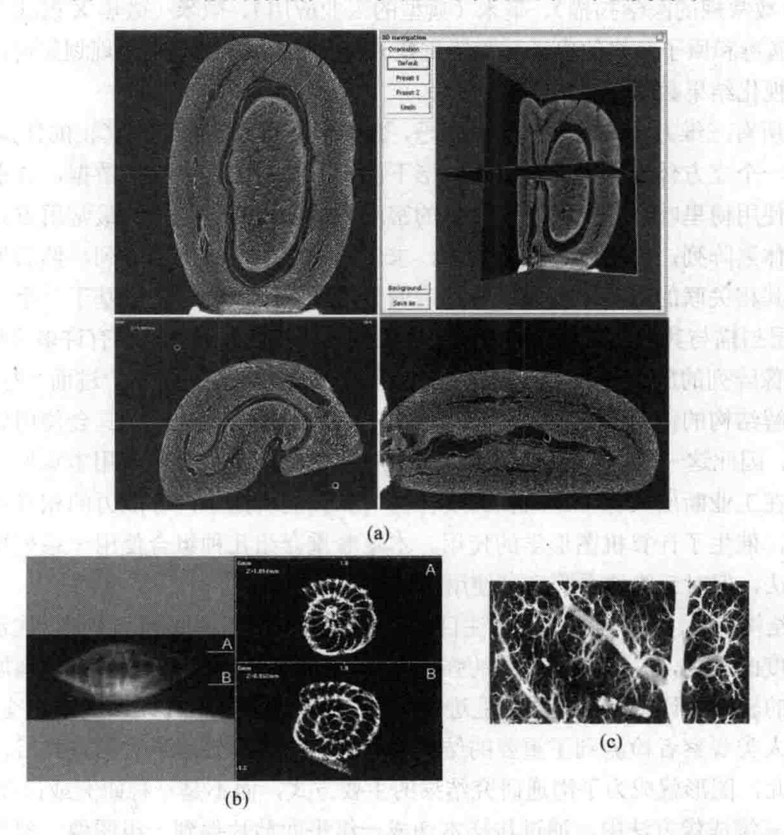


图 13.64　显微层析成像的几个例子：(a)通过咖啡豆的三个正交切片；(b)有孔虫的投影图像，它是使用通过标记位置的横向切片重建的；(c)老鼠肺部血管的三维可视化图像

第 14 章　三 维 视 图

14.1　三维数据源

生成三维图像数据的仪器（各种断层成像仪器、共焦显微镜等）的持续发展和计算机软件性能的提升，使得三维成像越来越容易。正如像素是二维图像的亮度测量单位那样，体素则是三维成像的单位。当像素是方形像素时，二维情形下的处理和分析很简单；而在三维情形下，首选立方体素，尽管实现立方体素通常并不容易。

三维成像有几种基本方法。第 13 章中介绍了通过层析重建的三维成像。这是测量密度和固体标本成分的最佳方法。它可产生一组立方体素，但这并不是断层成像目前所用的唯一方式或最为常见的方式。多数医学和工业应用会产生一系列二维切片，但这些二维切片的间距较大，或表示了一个比平面内横向分辨率更大的厚度（Baba et al.，1984，1988；Briarty & Jenkins，1984；Johnson & Capowski，1985；Kriete，1992）。

如前一章中所述，使用许多不同的信号可进行断层成像，包括地震波、超声波、磁共振、常规 X 射线、伽马射线、中子束、电子显微镜以及其他我们不熟悉的方法。分辨率可以是千米（地震层析成像）、厘米（多数常规的医学扫描）、毫米（典型的工业应用）、微米（微焦 X 射线或同步加速器射线源）甚至纳米（病毒和原子晶格的电子显微镜重建）。无论成像方式或维度规划如何，都可用相同的基本展示工具来可视化结果数据。

此处讨论的所有三维方法的一个重要区别是，数据集是像素平面还是真正的体素阵列。第 13 章中曾讲到可以建立一个立方体素阵列，从三维情形下的一系列视图收集投影数据，并求出（要么使用滤波反投影，要么使用傅里叶逆变换）每个体素的密度。但执行断层扫描的最常用方式是，一次将一个平面定义为一个体素阵列，收集一组线性视图，求出该切片的二维密度阵列，然后处理下一个平面。每个平面都有与其相关联的深度，因为这种技术的分辨率有限，但这些解基于一个二维模型。按这种方式使用时，断层扫描与其他二维成像方法（连续成像或连续切片技术）会有许多类似之处（和问题）。

观看此类图像阵列的放射科医生期望能在他们的头脑中合并这些图像，进而“看出”三维结构（放射科医生在已了解结构的情形下，通常会寻找他们不熟悉的那些结构，尤其会使用到一些识别疾病或损伤的特殊方法，因此这一过程非常有帮助）。越来越多的当代医疗系统使用本章所介绍的技术来直接展示三维视图。在工业断层成像中，结构的多样性（和预测所期望内容能力的相对不足）以及研究与解释的时间因素，催生了计算机图形学的使用。本章着重介绍几种组合使用一系列平行二维图像来显示三维结构的方法，但对三维体系阵列可使用相同的算法。

渲染的三维结构视图，其外观非常引人注目，且通常用于表达结构的重要细节。展示旋转（见图 14.1）或动态改变透明度的动画，是实现此目的的强大工具，但在印刷文档中无法很好地展示这些动画。多数生成此类图像的软件和计算机系统并不能进行实时处理，但有些工作站确实具有交互式旋转的能力。通常情形下，在人类观察者检测到了重要的结构信息并确定了有效展示它的方式后，就产生了最终的图形或动画。因此，图形就成为了沟通研究结果的主要方式，而不是一种研究或诊断工具。

在最常用的三维成像方法中，通过将样本切成一组平面截片得到一组图像，然后将这组图像堆积成体素堆。有时，切片过程是物理的。生物材料、织物甚至金属块，都可使用显微薄片机切成薄片，

然后对每个薄片成像（就像以正常方式观看各个切片那样）。收集并对齐图像就产生了三维数据集，数据集中的体素通常会在 z 方向拉得很长，因为与每个切片内的横向分辨率相比，这些切片要厚得多，且切片间的间距要大得多。

另一方面，二次离子质谱仪使用入射离子束每次从样本表面删除一层原子。入射离子束穿过质谱仪来从单个元素中选择原子，然后放大原子并在荧光屏上成像或由相机记录。由大量元素收集一组图像后，可产生样本的一个完整三维图。这种方法与切片成像的一个区别是，不存在对齐问题，因为在删除表面层时，样本块是固定的。另一方面，不同结构的刻蚀率可能不同，因此表面并不保持为平面，而这种起伏或差异刻蚀很难解释。在此类仪器中，体素深度可能非常小（本质上是原子大小），而横向分辨率与其相比则要大很多倍。

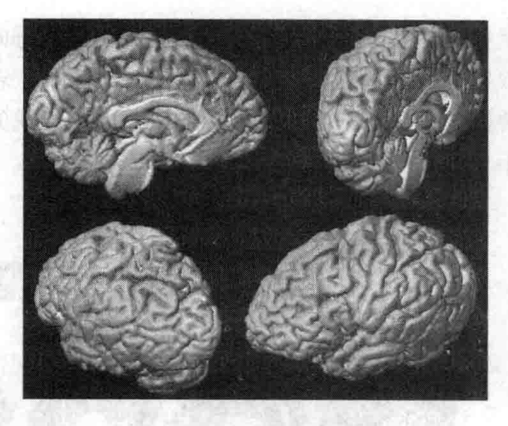


图 14.1　从动画中选取的几个静态帧，其中半个大脑的重建视图是不断旋转的

14.2　连续切片

多数物理切片方法都与这些例子相似。这些方法统称为连续切片法。尽管它对许多不同的情形都适用（例如在考古遗址按序剥除物质），但连续切片这一名称则源于使用光学显微镜对生物组织成像，此时使用显微薄片机将树脂中的组织块切成了一组切片。收集这些切片并在显微镜下观察，研究者就能收集一组可用于重建该三维结构的照片。收集图像并生成图形时，可使用许多商用和免费软件工具（Fiala，2005）。

使用基于一组切片的任何三维成像方法时，遇到的多数问题都可由连续切片技术来演示。如果已对删除每个切片揭示的表面成像，那么在对齐序列图像时就只有较小的失真。但如果分别对切片成像，则首先须对齐各幅图像。切割后的切片被收集到幻灯片或网格上，并从不同的旋转方向进行观察。因此，即使可在不同的截面中找到相同的结构（此任务并不轻松，必须假定结构的深度存在一定的变化），这些图片也不会对齐。

使用每个截面中可见结构的细节，只能进行大致对齐。各种自动对齐方法通常会使各截面间的失配最小，方法是对齐各平面中特征的质心，以便距离的平方和最小，或叠加来自两个截面的二值图像，并移动或旋转，以最小化这两个截面的“异或”运算导致的面积，详见第 8 章。这一过程如图 14.2 所示。图像中出现灰度值时，可以使用互相关法，详见第 6 章中的说明。遗憾的是，当截面在 X 方向和 Y 方向平移、拉伸或压缩而导致局部变形以及以任意角度旋转时，哪种方法都不容易实现。求解“最佳对齐”比较困难，通常必须反复地进行，因此速度较慢（Brown，1992；Averbuch & Keller，2002）。

此外，不要指望这些算法能到达最优点而真实地展现对齐。如图 14.3 和图 14.4 所示，平移或旋转每幅图像来目视对齐两个截面中的各个结构，完全可以改变重建的三维结构。在假定各幅图像中存在足够的细节时，使用某种类型的平均对齐可避免这些主要误差。但我们既不能确定最好的视觉对齐就是正确的对齐，也不能确定重叠序列图像的自动方法就能产生正确的对齐。

使用图像细节来进行对齐的一种改进方法是，在切片之前合并基准标记。例如，标记可以是激光钻孔、树脂变硬前放置的细线或纤维，或沿块边缘向下刻出的沟槽。对于某些（先对表面成像，然后向下抛光揭示另一个表面所得到的）不透明材料，表面缺口或划痕可用于对齐。使用在不同截面间形

状保持不变并在穿过堆叠图像的某个已知方向延伸的几个基准标记，可改善对齐。在感兴趣结构附近放置并找到基准标记通常很困难。实际中，若切片不连续，则也会出现问题，且对齐误差可能会在堆叠图像中传播。切割可能会导致截面中出现失真。这些因素都增大了对齐的困难性，特别是这些因素不一致且随位置变化时。

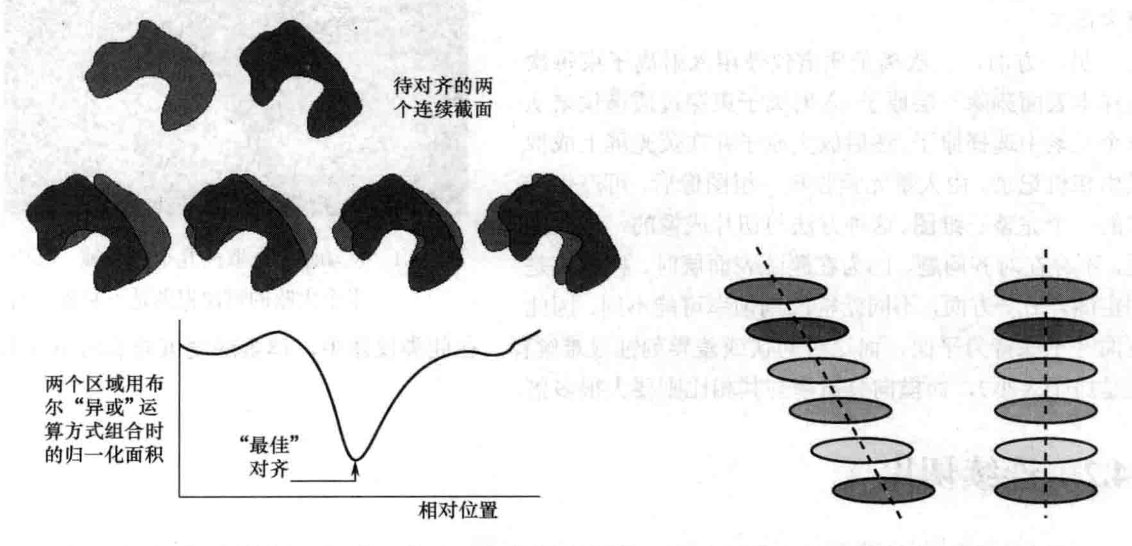


图 14.2 连续截面的对齐，"最佳拟合"最小化了特征区域的失配，特征区域是由"异或"运算测量的，且与平移和旋转相关

图 14.3 平移对齐连续截面：通过倾斜圆柱体的截面会被误认为是垂直的椭圆柱体

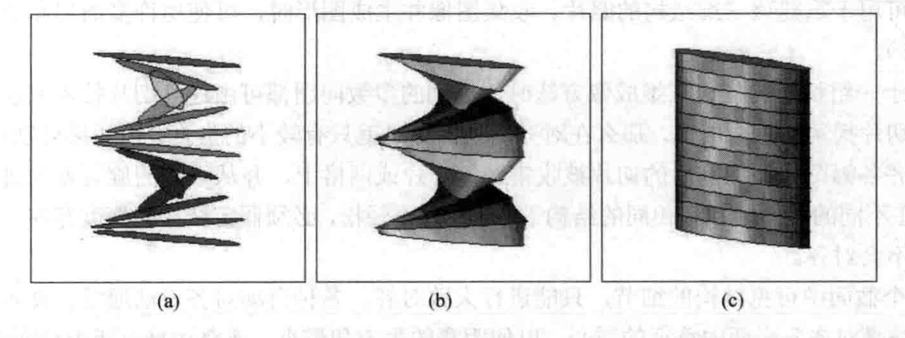


图 14.4 旋转对齐连续截面：(a)三维连续截面堆的实际轮廓；(b)轮廓的表面建模，显示了扭曲的结构；(c)轮廓彼此对齐后，消除了扭曲的错误结果

多数基准标记都会大到足以覆盖每幅图像中的几个像素。如第 10 章所示，这一尺寸允许我们以小数像素来定位质心。识别出对齐点（要么来自基准标记，要么来自内部图像细节）后，就可旋转和平移一幅图像，以将它与另一幅图像对齐，如第 4 章所示。对像素阵列进行重采样并内插来防止混叠，可生成一幅新图像。这个过程需要花费一定的计算时间，但与首次获得这些图像相比，计算时间要短得多。

遗憾的是，对于经典连续切片，旋转和平移的结果仍然不能真正表示原始的三维结构。使用显微薄片机切片时，通常会在块内产生一些变形。一个方向上的这种 5%～20%的压缩，通常假定对所有截面都是相同的（因为切片是按同一方向切割的，且结构上通常只有很小的差别，这种差别不会改变它们的机械性质）。如果已知基准标记的绝对坐标，那么可以通过拉伸图像来校正变形。通常假设整个切片是均匀压缩的，尽管对某些样品而言这一假设可能不正确。

如果三维重建的唯一目的是观看而非测量，则无须重点考虑结构与失真。在某些情形下，可使用内部信息来估计失真。例如，如果没有理由期望细胞或细胞核在组织中的任何首选方向被拉长，那么可使用许多细胞或细胞核的尺寸测量值来求压缩的平均量。显然，这种方法含有一些假设，且只能用在某些特定情形下。

连续切片的另一个问题是深度方向上的尺寸校准。各个截面的厚度只是近似已知的（例如，由顶部和底部表面的干涉产生的光的颜色进行判断，或基于切片机的机械进给率来估计，此时通常会忽略被切割材料的任何顺应性）。厚度可能会因截面不同而不同，甚至在截面内部的不同位置也不相同，具体取决于正被切割材料的局部硬度。很难构建准确的深度尺度，且深度方向维度的准确性要远小于某个剖面内测量的维度。

如果只用到部分截面，如每隔 1 个或每隔 4 个（为降低对它们成像并对齐图像的工作量），则该误差会增大。这也会使得人们逐幅图像来跟踪结构变得更为困难。但为降低必须处理的数据量，这种截面的跳过是必要的。

当切割的超薄截面在电子显微镜而非光学显微镜下观看时，只使用一小部分截面尤其常见。随着截面变薄，截面的数量会增加，因此更容易失真。有些截面可能会丢失（如由于折叠）或刻意跳过。每个截面中的一部分被支撑网格托着，因此无法用到。一种后果是，所获取图像的间距可能不均匀。在更高的放大率下，基准标记变得更大，因此定义就不那么精确，且分离得更开而不可能靠近感兴趣的结构。

图 14.5 显示了一系列组织的 TEM 图像的几个例子，其中我们感兴趣的是细胞膜的三维结构（深染色线）。细胞和细胞器的边缘细节已被用来近似地对齐通过堆叠的两个截面，但对不同的两个截面必须使用不同的细节，因此不存在连续通过整个堆叠的细节。对这些图像进行阈值处理后，可隔离出细胞膜（见图 14.6），但由于这些截面的间距太大，不能将这些线条连在一起来重建表面的三维形状。这一问题在传统连续切片图像中很常见。

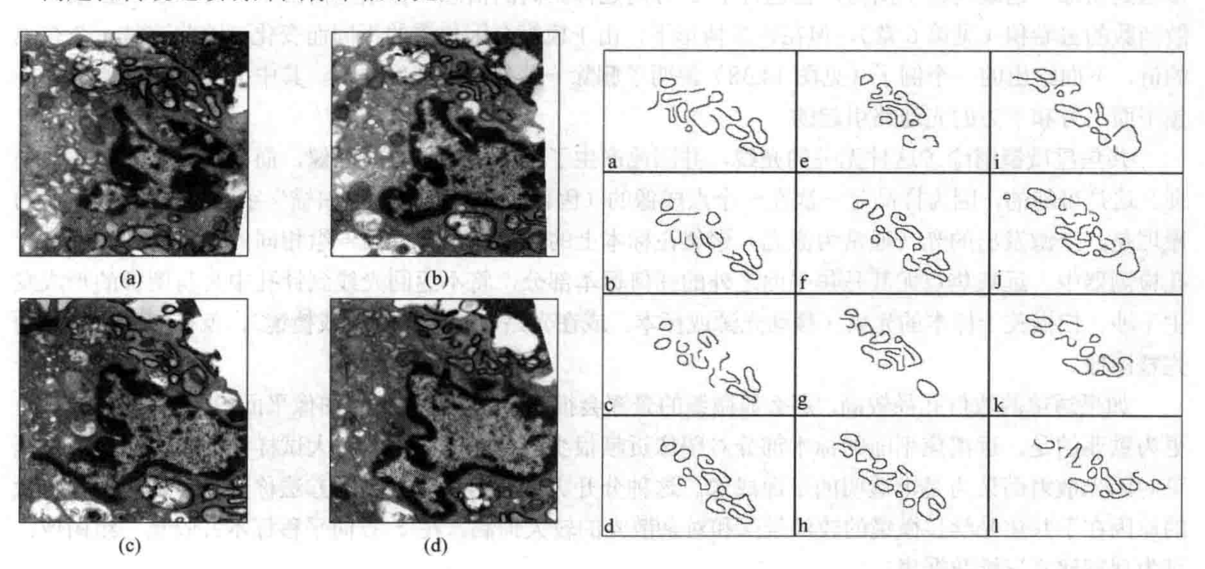


图 14.5 来自一个堆叠的 4 幅连续切片图像，它们已旋转并对齐。对图 14.6 中的整堆图像，图像顶部的细胞膜已被阈值分割并显示

图 14.6 图 14.5 所示序列图像的细胞膜，显示了截面的变化。由于截面间的间距原因，这些变化太大，因而不能重建表面的三维形状

对于不透明材料的反射光成像，使用物理抛光方式顺序地移除材料层，或使用聚焦离子束来进行腐蚀，可实现连续切片。如果成像时样品的位置固定，那么找到同一位置并监控抛光的深度就很简单。

如果样品未固定，那么表面刻痕或激光蚀坑就可作为对齐之用的基准标记。在考古挖掘中，基准标记可能是绳网和经纬仪，而清除工具可能是泥刀或铲子。有些采矿及采石的例子可能会使用经纬仪来对齐，而使用推土机进行清除，但无论规模大小，其原理都相同。从表面顺序移除材料将在下面讨论。

在多数表面成像方法中，这些连续的二维图形表示在 z（深度）方向上分开的平面上的样本。必须推断出或从两个平面内插出已被清除的中间材料。此时，像素值不是整个距离上的平均。相距较远（根据结构的规模）的两个截面间的内插，可能会导致一些严重的错误和误解。

对于以透射方式查看切片的连续切片方法，像素值是穿过厚度的体积平均，且可当做一个体素来处理，因此解释起来很容易。在多数情形下，体素值是材料密度的测度。取决于使用哪一种辐射（可见光、X 射线、电子、中子、声波等），该值可表示某些元素或化合物的局部浓度，包括作为示踪剂或标记的那些化合物。在有些情形下，从体素发出的射线也可给出浓度信息（例子包括荧光显微镜和正电子发射断层扫描）。

14.3 光学切片

任何尺度的物理切片都会破坏样品。控制并测量截面厚度和对齐截面，或至少找到样品上的同一位置，可能是主要的误差来源。在某些情形下，无须执行物理切片就可以对穿过样品的截面成像。共焦扫描光学显微镜（CSLM）提供了一种方式来实现此目的（前一章中介绍的切片图像的层析成像重建是另一种方式）。在某些情形下，通过解释在光学显微镜下观察物体引入的相移，以及由中子或 X 射线透射成像引入的相移，也可获得深度信息（Barty et al.，2000）。

正常运作的透射光显微镜本身并不进行光学切片。高数值孔径光学系统的景深很小（是横向分辨率的几倍），所以图像中只有一个小“切片”是清晰聚焦的。但焦平面上方和下方位置的光也会失焦地传送到图像，这既模糊了图像，也包含了 z 方向延伸距离的信息。在某些情形下，可以实现三维点扩散函数的去卷积（见第 6 章），但在许多情形下，由于模糊会因位置的不同而变化，因此这是一个有限的值。下面给出的一个例子（见图 14.38）说明了删除一些伪影的图像处理，其中的伪影是由通过样品焦平面上方和下方的光通路引起的。

共焦显微镜消除了这种无关的光线，并因此产生了有用的光学截面图像，而不需要对图像进行处理。这是可能的，因为样品是一次在一个点成像的（因此名称中出现了“扫描”字样）。共焦显微镜的原理是，点源发出的光（通常为激光）聚集在标本上的一个点处，并由一组相同的光学器件收集到针孔检测器中。远离焦点尤其是焦平面之外的任何标本部分，将不返回光线到针孔中来与图像的形成发生干涉。扫描关于标本的光束（移动光源或标本，或在光路中使用扫描镜或棱镜），就可得到焦平面的完整图像。

如果透镜的数值孔径较高，那么显微镜的景深会很小，但它仍比各个图像平面内的横向分辨率大。更为重要的是，远离焦平面的标本部分对图像贡献很少。这使得有可能对大试样中的平面成像，甚至可对因光散射而认为是半透明的平面成像。这种分开大试样内单个平面的方法称为光学切片，其有效的原因在于共焦光学显微镜的较浅景深和对杂散光的较大抑制。在 z 方向平移标本并收集一组图像，可为观察建立三维数据集。

共焦光学显微镜有几种成像方式。最常见的是反射光，此时从样品反射的光会从对入射光起聚集作用的同一物镜返回，然后被平面镜反射到检测器；另一种是荧光，此时的光由样品内的各点发射，并使用相同的几何排列进行记录。很少使用显微镜来观察透射光图像（图 14.7 显示了一个透射光焦平面截面）。

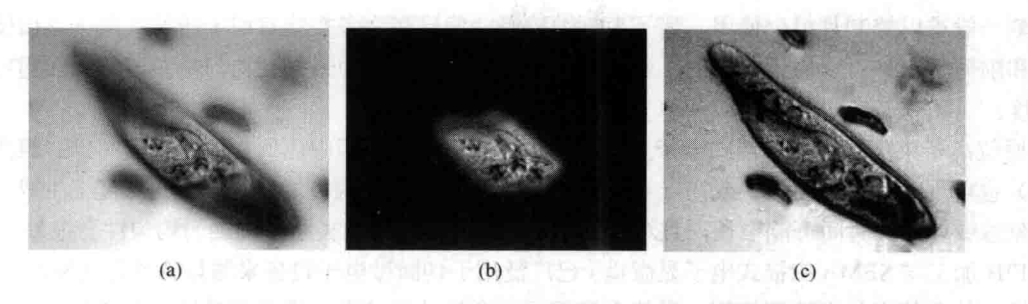


图 14.7　水滴中游动的草履虫图像：(a)传统的透射光显微图像；(b)透射共焦显微图像（只有部分焦平面可见）；(c)合并来自共焦系列的所有平面，所产生的扩展聚焦图像

焦平面截面图像可用于不同类型的标本的三维成像。反射光共焦图像的特点是，反射回检测器的光强会随点在焦平面上下移动时快速下降。因此，对于透明介质中的结构，只有表面反射光。对任何一个图像平面，只有某些结构穿过该平面的部分视野看起来是明亮的，而剩下的视野则是暗的。这一特性允许我们直接重建算法。

在发射成像或荧光成像中，入射光的波长能够激励引入到标本中的染料或其他荧光探头。这种探头发射的低能量（较长波长）的光与入射光分离（如由一个分色镜分离），并用于形成一幅图像，在这幅图像中，探针或染色位置看上去是明亮的（第 8 章的图 8.29 给出了一个例子）。按深度建立一系列图像，就可由探头标注将被重建的结构。

光学切片的主要优点是速度快和易于使用，可避免物理切割引起的标本变形，保持来自各个成像平面的图像的对齐，直接控制切片的 z 方向位置。深度分辨率尽管比每个平面的横向分辨率差 2～3 个因子，但对许多应用它仍是有用的。但这种分辨率的差别确实会增大三维图像处理的难度，即使平面间距受控以便存储的体素是立方体素。

通过对显微镜的点扩散函数的三维形状进行测量或建模，可通过去卷积操作来提升共焦光学显微镜的分辨率。该方法与第 6 章所示的用于二维图像的方法相同，但是以三维傅里叶变换执行的（傅里叶变换是可分解的，这意味着它可沿像素或体素的行和列逐个并顺序地执行）。

14.4　连续移除

许多材料是不透明的，且不能通过光学切片成像。金属、复合材料和陶瓷通常在反射光显微镜下检查。但如上所述，通过连续抛光这些材料，仍可收集到一系列深度图像来进行三维重建。

从表面移除材料的方式，很大程度上取决于材料的硬度。对于一些软金属、聚合物和织物而言，就像对生物材料使用薄片切片机那样，也可使用薄片切片机，只是此时不检查已移除材料的切片，而是对剩下的表面成像。这种方法避免了多数对齐和失真问题，尤其是切割在原位进行而不从显微镜下的观察位置移动标本时。精确地求出每次切割所移除的材料的厚度并保证其均匀性，仍然很难，但由设备上的机械设置通常可估计出材料的厚度（忽略材料中的任何永久失真或临时失真）。

对于较硬的材料，可以使用用来对二维图像产生传统截面的研磨或抛光技术。此类操作通常需要移除和替换标本，因此需要用基准标记来定位同一区域。进行这种标记的最普通方法是使用硬压痕。在标本的表面形成几个金字塔形的压痕，以便在磨损或抛光后，压痕最深的部分仍然可见。这些压痕可与原始图像中的标记对齐。此外，减小已知形状压痕的尺寸，可测量出抛光深度以及两幅图像间的间距。使用几个这样的压痕，也可判断抛光的整体均匀性，尽管会出现特殊相的硬度引起的局部变化。

对于更硬的材料，或传统抛光可能会导致表面损伤的材料，可以使用一些其他的方法。电解或化

学蚀刻一般难以控制且很少使用。离子束刻蚀较慢，但已在许多实验室用于透射式电子显微镜标本的切割和削薄，也许可以用于此目的。对许多标本而言，控制刻蚀来得到均匀性并避免表面粗糙有较大的难度。

原位离子束刻蚀用于扫描式电子显微镜和扫描式俄歇显微镜中，例如清除表面污染。聚焦离子束（FIB）也可用于移除表面材料或在 z 方向上切片来研究微观结构。这种能力可在深度上生成一系列图像。刻蚀或切片的时间可能很长，且对复杂结构（最有趣的成像类型）刻蚀的均匀性会很差。

FIB 加工与 SEM（扫描式电子显微镜）已广泛用于切断微电子设备来测量临界厚度和尺寸。此时，根据标本的几何形状来控制切割，对能直接提示重要尺寸的结果，通常会进行垂直切割。当然，也可使用 FIB 加工来产生可进行体视测量的截面，或将一系列切片用做连续切片装备来直接揭示三维结构。对金属和陶瓷微观结构的测量，已有限度地使用了这种方法（Holzer et al.，2004，2006，2007）。这类机械加工所揭示的表面堆叠图像可用于三维重建，包括尺寸的测量和结构元素的拓扑。图 14.8 显示了一个例子，例中测量了颗粒并根据大小对颗粒进行了着色，对颗粒间的相交面也进行了着色（Holzer，2006）。使用 FIB 技术也可对一些组织标本切片。但这些图像的主要用途一直是视觉显示而非测量。

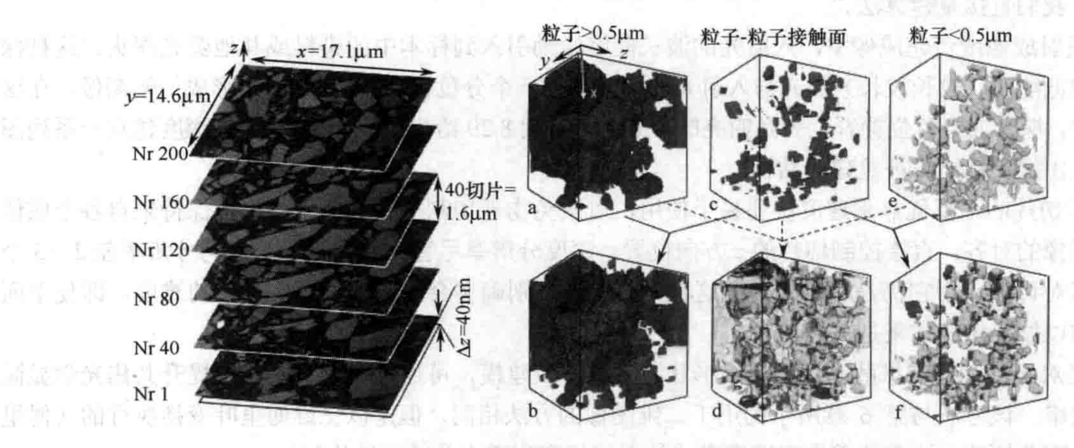


图 14.8 使用 FIB 表面加工技术得到的连续截面堆叠图像，以及对不同大小颗粒和各个相交面使用不同颜色重建的效果

离子显微镜或二次离子质谱仪（SIMS）会自动地刻蚀标本表面，这一自动操作是其图像处理过程的一部分，即利用重离子束来从标本表面刻蚀一层原子。然后根据质谱仪中的元素分开二次离子并进行记录（例如使用一个通道板倍增器和数码相机），进而每次一个元素地形成标本中一个平面的一幅图像。对于这些仪器，通过测量一种已知标准的信号剖面，可校准刻蚀的深度。

表面去除速率是高度可控的，且在深度上基本上是原子分辨率。相比之下，横向分辨率大致与常规光学显微镜的分辨率相同。此外，离子显微镜或 SIMS 中离子束撞击的刻蚀速率，会因标本中位置的不同而变化，它与成分、结构或晶向有关。因为假定每组数据表示了一个平面，因此重建结果中无须显示这一变化，但在最终解释时会导致明显的失真。原理上，就像校正二维图像的变形那样，可对三维数据进行拉伸。但如果局部刻蚀速率没有基准标记或准确的量化数据，就很难使用这些真实的数据进行重建。

使用 SIMS 对许多不同元素成像的能力，为三维显示创建了丰富的数据集。彩色二维图像有 3 个通道（见第 1 章），但 SIMS 数据实际上可能存在任意数量的通道，分别表示元素、同位素和分子碎片。设备检测这些微量元素的能力，意味着即使对于相对简单的标本而言，也难以存储、显示和解释多波段数据。

对原子层成像时，移除原子层的另一种显微镜是原子探针离子显微镜。在这种仪器中，急剧弯曲的样本尖端和显示屏之间的强大电场，会使表面上的原子解除吸附，并加速撞向它们被成像的屏幕。

屏幕可包含一个电子通道板来放大信号，以便可看到各个原子，或这些原子可用做高压脉冲飞行时间质谱仪，进而区分不同的原子。随着仪器的变化，结果是样本原子的一幅高倍放大图像，随着一层一层地移除原子，就揭示了原子的三维排列。许多此类仪器的图像例子请参考第 1 章。

14.5 立体测量

使用另一种方法也可看到三维结构，即人类在某些真实情形下感知深度的方法。两眼朝前并重叠视野，就可使用立体视觉来判断物体的相对距离。如第 2 章所示，人类是通过在眼眶中移动眼球，使每个物体都投射到中央凹（视网膜中视锥细胞最密集的部分），来逐点这样做的。肌肉反过来告诉大脑需要什么动作来实现会聚，进而指出哪个物体更近，哪个物体更远。立体视觉在下面也用做向人类观察者展示三维数据的方式。

所有人类的深度感知都依赖于体视学这种想法是错误的。事实上，我们关于三维世界的许多判断均来自其他线索，如阴影、相对大小、先后顺序、大气影响（如雾或霾）和相对运动（随着头部的移动，较近的物体在视野中移动得更快），例如在某些基于计算机的测量方法中就使用了这些线索（Roberts，1965；Horn，1970，1975；Woodham，1978；Carlsen，1985；Pentland，1986）。但使用立体图像可确定深度信息，并将信息放到三维计算机数据库中。

光学显微镜的景深较浅，因此研究起伏较大的标本时，结果很不让人满意，除非是在低放大率下。但是，电子显微镜的透镜具有非常小的孔径角，因此具有非常大的景深。结合使用体视学与扫描电子显微镜（SEM），可产生粗糙表面的图像。平铺标本或等效地偏转扫描光束，可以由形成一个立体像对的不同视点生成一对图像。用两只眼睛注视一张图片，可欺骗大脑感知原始的粗糙表面。

由此类图像来测量起伏表面的原理，与使用航拍或卫星立体像对来测量地球或其他星球的地形要素高度的原理相同。卫星照片上的丰富细节，使得我们可很容易地在图像中的任何位置找到匹配点，但出于同样的原因，与在 SEM 中观察到的简单标本几何相比，需要更多的匹配点来定义表面。测量视差之间（左、右眼图像中点的偏移距离）的数学关系和表面上各点的相对高度，已在第 1 章中介绍。

由立体像对自动匹配各点，对于基于计算机的图像分析而言是一项艰巨的任务（Marr & Poggio，1976；Grimson，1981；Medioni & Nevatia，1985；Kayaalp & Jain，1987；Smith & Elstrom，2001）。常见的匹配过程如下：使用一幅图像（如左侧图像）中的亮度值模式作为模板来执行互相关操作，搜索右侧图像中最接近相同的模式。搜索区域被限制为第二幅图像中覆盖可能偏移的水平条带，具体取决于两个视图间的角度以及表面的最大粗糙度。这一过程不会匹配某些点，因为这些点在两幅图像中均不可见，而有些点的匹配效果很差，因为像素中的局部亮度值模式含有噪声。当视点变化时，在表面上可能移动的镜面反射也会干扰位置的匹配。

匹配许多点会产生一幅新图像，其中的每个像素可根据侧移来赋一个值，进而表示表面的高程。这幅距离图像可能包含许多伪匹配，但中值滤波这样的运算通常可很好地删除异常点，产生表面的整体距离图像。在图 14.9 所示的例子中，左眼视图中的每个点与右眼视图中的每个点的互相关匹配，产生了一幅含有伪匹配的视差图（匹配点位置间的水平距离），但它已被中值滤波消除，如图所示。如图中说明的那样，生成的高度数据可用于测量点或线轮廓，或用于重建表面图像。表面距离数据的这种用途将在第 15 章中详细探讨。

匹配立体像对的第二种方法则基于这样一种现实，即每幅图像中的许多点都不能很好地匹配，因为它们不是表面相交或出现其他中断的“兴趣”点。这种方法或许与人类视觉有关，它会花许多时间来集中精力于在每个场景中可找到中断的那些少数点上。根据一些局部性质（如方差、熵或高通滤波的结果）来找到这些感兴趣点，可生成在两幅图像间匹配的一个相对较小的点表（Moravec，1977；Quam & Hannah，1974）。典型情形下可能只有几千个点，而不是原始图像中上亿个像素。

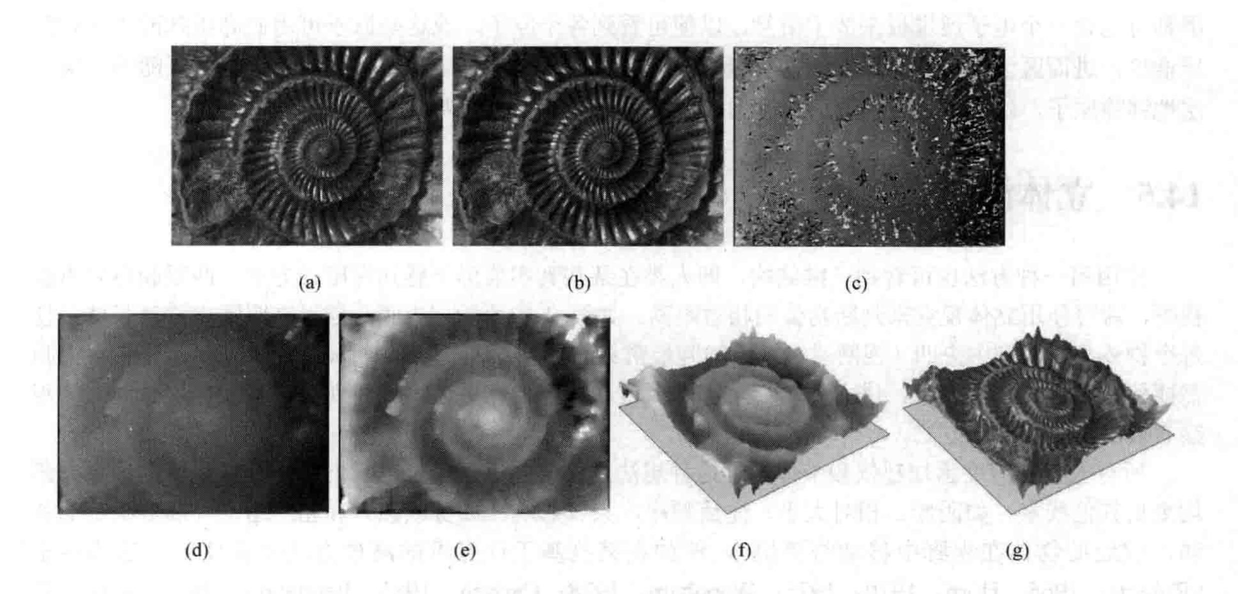


图 14.9 使用立体像对测量高程：(a)微化石的左眼视图；(b)微化石的右眼视图；(c)原始的互相关视差值；(d)图(c)经中值滤波后的结果；(e)从一个平均平面测量的表面高度值，显示为灰度值；(f)使用(e)中的值渲染的透视表面模型；(g)使用(e)中的高程值及(a)中的表面亮度值重建的同一表面

将一个三次多项式拟合到每个邻域中的像素亮度值，可更好地选择这些点。因此，如果该多项式写为

$$f(x,y)=c_1+c_2x+c_3y+c_4x^2+c_5xy+c_6y^2+c_7x^3+c_8x^2y+c_9xy^2+c_{10}y^3 \tag{14.1}$$

则用于检测角点的 Zuniga-Haralick 算子（Zuniga & Haralick，1983；Haralick & Shapiro，1992）为

$$\frac{-2(c_2^2c_6-c_2c_3c_5-c_3^2c_4)}{(c_2^2+c_3^2)^{3/2}} \tag{14.2}$$

所产生的较小点表上的点按上述方式匹配，方法是与它们的邻域亮度模式相关。此外，对于多数表面（即没有环或桥的简单连通表面），按从上到下、从左到右的顺序保留这些点。它和某个给定像素对的视差限制，一起将典型候选点表中的点数减少为一个很小的数字（通常为 10 或更小），进而产生一序列表面点和它们的高程。然后假定这些点间的表面表现良好，并可认为它由小平面或简单的样条块组成，而在 Delaunary 三角网中连接这些点，就可形成这些小平面或样条块。如果小平面足够小，那么可生成表面的等值线图，如图 14.10 所示，方法是在每个小平面的边缘出现的中间，内插直线段或样条曲线。称为距离图像的完整高程显示可通过内插产生，如图 14.11 所示。

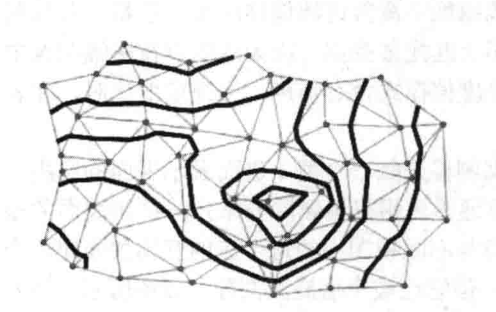


图 14.10 任意连接其高程已通过体观学确定的点后，三角面上显示的等值线（等高线）图

透射电子显微镜（TEM）的景深非常大。在多数情形下，在 TEM 中观察的标本非常薄（以便允许电子渗透），因此光学景深并不重要。但在使用高压显微镜时，可对相对较厚的样本（微米级）成像。这一厚度足以在 TEM 的分辨率中（几个纳米级）包含大量的三维结构。采用倾斜标本的相同方法来获取立体像对时，也可获得关于点的深度和三维结构的信息。

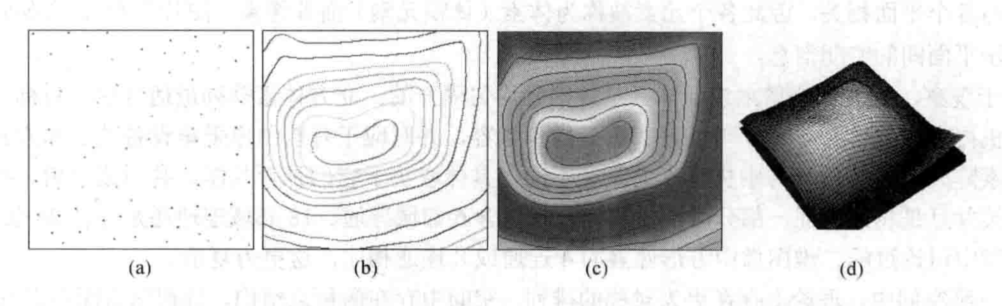


图 14.11　距离图像的内插：(a)孤立且随机分布的高程点（已着色）；(b)由曲面细分画出的等值线；(c)等值线间平滑内插后的结果；(d)构建的表面几何

人眼在观看两幅图像时，会看到场景中物体的距离并不相同，因此在脑中融合于不同视点获取的两幅图像并进行深入检测并不困难（第 1 章中给出了几个例子）。这项技术似乎每隔 10 年时间就会重新流行一次，既用于静态图像，也用于电影，也许很快就会用于电视。在许多情形下，它们都是表面图像，而不是透明体中的物体。

自动匹配点和测量深度的部分问题是，会出现围绕感兴趣特征的透明介质的背景细节。这种问题可能会支配局部像素值，并使得匹配不可能。另一个问题是，不能再假设这些点的顺序是从左到右。在三维结构中，当这些点彼此前后穿过时，就会改变顺序。

这些限制的后果是，仅在少数理想情形下成功地进行了 TEM 立体像素的自动融合。使用对比度非常高的标记（如与使用抗体选取的表面绑定在一起的小金粒，或某些其他高度选择性的染色剂）来简化这一问题是有帮助的。此时只考虑这些标记。仅有几十个或至多几百个这样的标记，且像上面介绍地图表面时提及的感兴趣点一样，很容易检测出这些标记（它们通常要远暗于图像的其他部分），但只有少数标记匹配。

即使使用这些标记，也仍需要人来识别或编辑这些匹配。给定两幅图像中的匹配点后，计算机就可以构造出表示由标记定义的表面的一组直线，但这个表面可能只是整个结构的一小部分。图 14.12 显示了人工匹配点的一个例子（Peachey & Heath，1989）。类似的方法可应用到染色网络（Huang et al.，1994）或材料中沉淀微粒的分布。

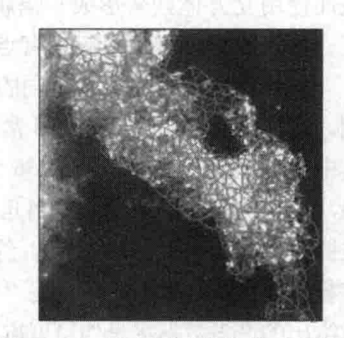


图 14.12　用金属颗粒装饰表面的例子（高尔基染色法）。以立体方式确定的网格叠加到了透射电子显微图上

体视学求出的高度误差通常要比视差测量精度大一个数量级，因为垂直高度与水平视差和视图间较小夹角余割的乘积成正比。因此，应将特征间的视差测量精度提高到亚像素精度。这种改善在某些情形下是完全可能的，特别是当来自许多像素的信息可以组合时。如第 10 章所示，可为特征的质心或直线的位置指定小数像素的精度。

14.6　三维数据集

存储三维数据集的常见方式是存储一组二维图像。之前描述为像素阵列的每幅图像，现在可理解为拥有或表示深度。出现这一深度的原因，要么是该平面确实是样本的某个深度之上的平均值（如透过一个薄截面），要么基于该平面与下一个平面间的间距（如通过反射光观察到的一组抛光平面）。由

于深度与各个平面相关，因此各个元素被称为体素（体积元素）而非像素（图片元素），即使它们并不完全表示平面间的空间信息。

对于观察、处理和测量来说，体素是规则且均匀隔开的。立方体素阵列最适合这一目标，因为它在计算机内存中最容易寻址，与处理和测量操作兼容，并对应于有些图像采集设备的工作方式。采用立方体素阵列，相邻体素与中央体素的距离不同，具体取决于它们是否共面、共边或共角。决定体素是否相接并且是相同特征一部分时，需要确定是包含 6 邻域连通、18 邻域连通还是包含 26 邻域连通，与第 7 章中讨论过的二维图像中方形像素的 4 连通或 8 连通相比，这更为复杂。

在三维空间中，理论上存在更为对称的排列。空间中存在两种点结构，分别称为面心立方（FCC）晶格或六方密积（HCP）晶格，这种晶格中，每个点（或体素）周围排列了 12 个等距的邻点。这些体素堆叠排列的好处是，可以等同地对每个邻点进行图像处理，且测量值基本与方向无关。邻点距离更均匀的一种更对称的邻域，也可简化处理过程（包括应用滤波和形态学操作）。但为了填充空间，这些情形下体素的形状是棘手的多面体，对体素进行存储和寻址很困难，获取图像或显示图像也很困难。通常情形下，所获取的图像必须通过内插来重新采样，以便以其中的一种模式来得到体素值，显示时则需要进行反内插。对于多数目的，这些不利因素的影响超过了理论上的优势。正如大部分二维图像使用方形像素阵列而非理论上更具吸引力的六边形像素阵列那样，立方阵列是最常用的体素三维排列。

如果由于平面间的间距不同于每个平面内的分辨率导致体素不是立方体素，则可进行调整以便体素是立方体素。以下的讨论假设深度间距大于平面内的间距，也可为相反的情形描述一种类似的情况。在已测量体素平面间内插出其他体素平面，可完成这一调整。遗憾的是，这样做对图像处理操作并无多大帮助，因为假设是所有的相邻体素都是同等重要的，而内插会使得它们有点多余。

另一种方法是，通过对每第 n 个像素采样，或对块内的像素进行平均，可降低平面内的分辨率，进而使用立方体素来形成一幅新图像。这种分辨率的降低可减少所需的存储量，因为保留的体素更少了。尽管放弃分辨率看起来有点反常，但在要求使用立方体素进行分析的情形下，确实可以这样做。

三维数据集可以许多不同的格式来存储，不管是各个切片的堆叠，还是使用 x, y, z 索引的体素值阵列。这种阵列很快就会变得非常大。一幅 1024×1024 像素的二维单色 8 位图像要占用 1MB 的存储空间，其中每个像素占用 1 字节（256 个灰度值）。在台式计算机的内存中，这很容易处理。即使是新款数码相机，典型的彩色图像也只占用几十兆字节的存储空间。但 2 字节/体素的 1024×1024×1024 体素的一幅三维图像需要 2GB 的内存空间，约为台式计算机的实际处理上限（因为许多处理运算需要两个数组，且许多程序会保留一个副本来满足“恢复”的需要）。较大的文件难以存储和传输，处理起来就更为困难。本章给出的许多数据集是切割或取样自较大数组的一些小数组。给出的操作适用于较大的数据集。

将这种情形与计算机辅助（CAD）进行比较是有益的。对于具有相对简单几何表面的人造物体，只需极少数量的点坐标和尺寸来定义整个三维结构。这类边界表示非常紧凑，但往往需要花费一定的时间（或专门的显示硬件）来渲染出来自这样一个数据集的真实表面的图形。对于一幅体素图像，存储需求非常大，但可立即得到每个位置的信息，而且本章中给出的多数不同显示图像，可以由适当的计算机快速生成（假设数据正在内存中）。

例如，给定由边界表示或一些坐标定义的一组表面，那么首先为一个平面构建所有点，计算该平面关于观察者和光源的局部角度，使用它们求出一个亮度值，并在屏幕上画出该值，即可生成显示。同时，使用另一幅图像内存来存储该表面在每个点处的实际深度（z 值）。完成一个平面的处理后，接着按相同的方式处理下一个平面，但要将深度值与 z 缓冲区中的值逐点进行比较，以确定该平面是在先前值之前还是在先前值之后。如果在之前，则只需画出每个点。该过程允许在屏幕上正确地画出多个相交平面［关于三维 CAD 数据图形显示的详细信息，请参阅（Foley & Van Dam，1984）或（Hearn & Baker，1986）］。

需要使用额外的逻辑方法来修剪平面的边缘，进而存储边界，更改根据表面特性来计算亮度的反射率规则等。关于计算机图形学的标准教材给出了完成这些任务的算法，并用大量篇幅介绍了各种方法的相对效率，因所涉及的时间相当重要。通过比较，查找较大数组中的值，或甚至处理该数组中的一列数据来添加宽度或找到最大值，速度是非常快的。如果可将数组调入内存而不是进行磁盘存储，速度会更快。

对齐连续切片来产生三维数据集的难度如上所述。在许多情形下，由不同成像技术（如头部的MRI、CT和PET图像）得到的几个三维数据集，必须相互对齐。这些数据集通常具有不同的分辨率和体素大小，因此需要内插来调整它们，以便彼此匹配。这种情形类似于在地理信息系统（GIS）中遇到的二维问题，即表面图、来自不同卫星的不同波长的图像、航拍照片和其他资料必须对齐并合并。

常见的问题是多个数据集的配准。第6章中介绍的两种互补的主要技术，对整个像素阵列或体素阵列使用互相关方法，或在多幅图像中分离出特定的特征，并使用这些特征作为基准标记来完成扭曲操作（Brown，1992；Besl，1992；van den Elsen et al.，1993，1994，1995；Reddy & Chatterji，1996；Frederik et al.，1997；West et al.，1997；Hajnal et al.，2001；Modersitzki，2004；Goshtasby，2005）。

14.7 对数据集切片

由于多数三维图像数据集都存储为一组二维图像，因此很容易就可存取各个图像平面或切片。为创建动画或“电影”，回放这组切片通常可有效地让观众进行三维合成，且在概括图像获取方式的同时，多数观众都能理解以这种方式呈现的图像。例如，使用简单的用户界面就可让观众改变动画的速度、改变方向或在某个选取的切片上暂停。这是迄今为止交互地检查三维数据集的最常见方式。

以切片数据来呈现原始图像的一个问题是，三维结构中某些特征的方向在切片中可能显示得并不好。不管是静态回放还是动画回放，能够改变切片的方向来研究穿过数据的任何平面，都是很有用的。只要切片的方向平行于数据集中的 x、y 或 z 轴，就可很容易地改变方向。如果将深度方向视为 z 轴，那么 x 轴和 y 轴就是各幅图像的水平边缘和垂直边缘。如果数据存储为离散的体素，那么存取该数据来生成平行于这些方向的图像或平面，就是计算各个体素与数组中每行及每列开始处的偏移地址。如果数据已调入内存，那么几乎可以实时地完成这种寻址；但如果数据存储在磁盘驱动器上，则寻址速度会较慢，因为原始切片图像中沿扫描线相邻的体素在磁盘中是连续存储的，一次扫描只能读出一组体素。但在需要不同的方向时，则须使体素广泛分布在文件的不同位置，因此存取磁盘需要时间。

第13章中介绍了如何在平行于 x 轴、y 轴和 z 轴的各个平面上显示图像。图14.13给出了正交切片的另一个例子。这些图像是人体头部的磁共振图像（MRI）。各个视图通常是轴向的（垂直于人体的脊柱）、矢状的（平行于脊柱和主对称轴）和冠状的（平行于脊柱并垂直于视线的正前方）。图中显示了表示这些方向的几个截面。

对MRI数据或其他类型的医学图像重新切片，会使质量下降，因为平面间距通常要大于平面内的分辨率，且在重新切割的切片中的某个方向上，由于 z 方向内插的影响，分辨率下降明显。替代内插的一种方法是在空间中延长体素；多数情形下，这会更加分散眼睛的注意力，如图14.14所示。像素平面间的内插可以线性地完成，也可使用高阶多项式来拟合两个以上的平面。但是，尽管内插产生了视觉上可以接受的图像，但它会模糊真实的结构或创建可视结构。图14.15显示了内插的一个例子，该例子创建了一个实际上并不存在的结构。

在本章的几幅图中，通过对原始数据采样并内插来重建平面图像时，z 方向分辨率的下降非常明显。在图14.13的情形下，MRI图像是在三个方向均有一致分辨率时得到的。许多当前的诊断过程会获取不同方向的多组扫描数据，进而在任何主要方向提供高质量的观察截面。

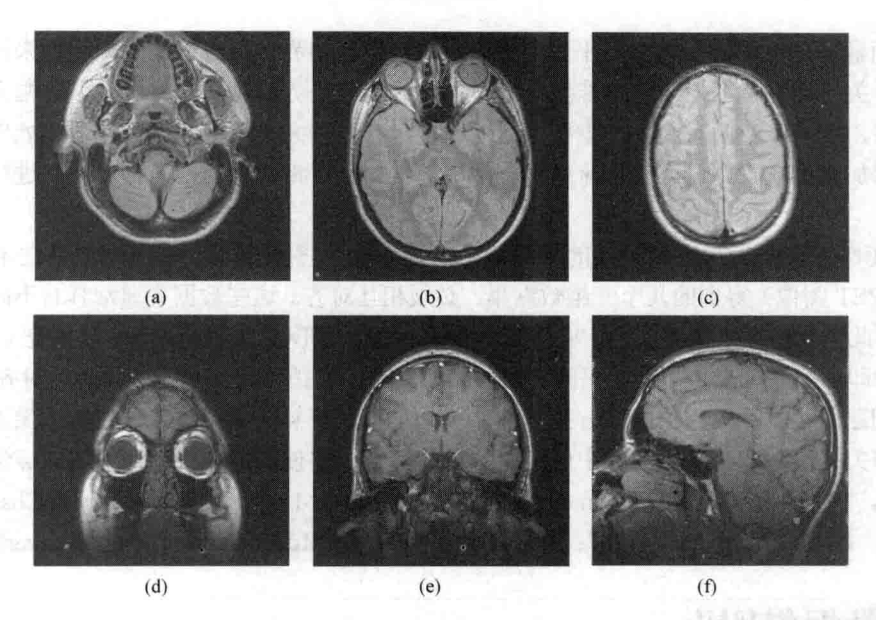


图 14.13　来自整套核磁共振成像（MRI）头部扫描数据的几个切片：(a)～(c)轴向切面（46 个切面中的 3 个切面）；(d)～(e)冠状切面（42 个切片中的 2 个切片）；(f)矢状切面（30 个切片中的 1 个切片）

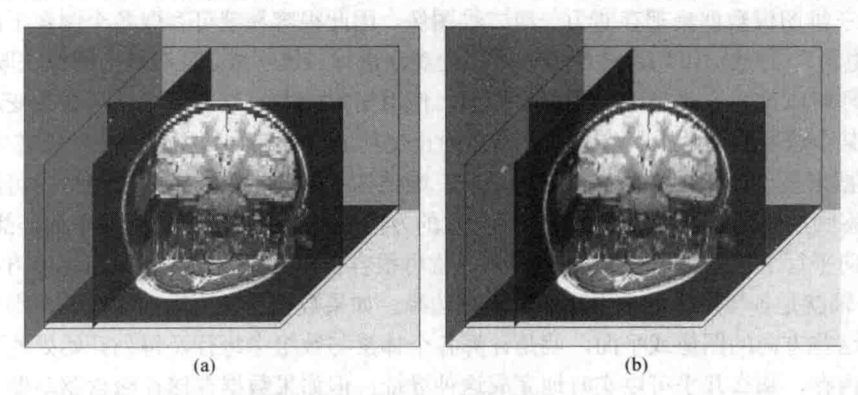
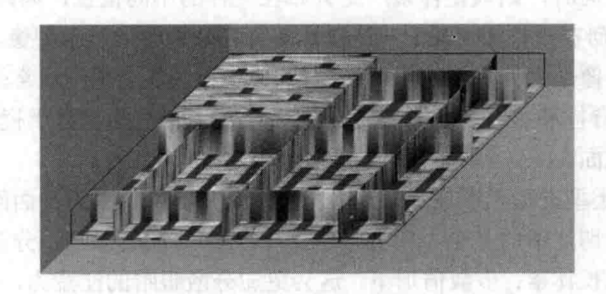
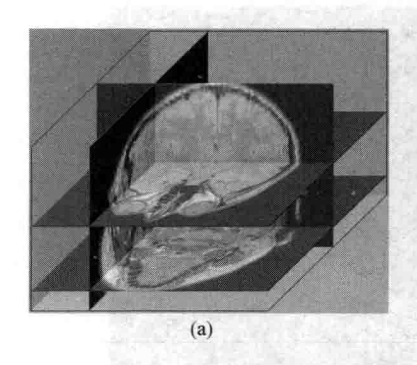


图 14.14　穿过图 14.13 所示三维 MRI 数据集的两个垂直切片的比较：(a)垂直延伸的切片；(b)两个切片间的线性内插


图 14.15　一个三维数组中的内插。该例中只使用了织物顶部和底部的两幅图像。两个表面之间的那些点是线性内插的，而与织物的实际三维结构无关

使用正交平面同时组合几幅视图可增加数据的三维效果。图 14.16 显示了几个这样的例子，例中使用了相同的 MRI 数据，该数据是基于轴向的平面图像得到的。*z* 方向的分辨率较差，但三维结构的总体效果仍然较好。移动穿过数据集的一个（或几个）平面，同时固定其他正交平面作为视觉参考，就可由这些视图生成动画。


(a)

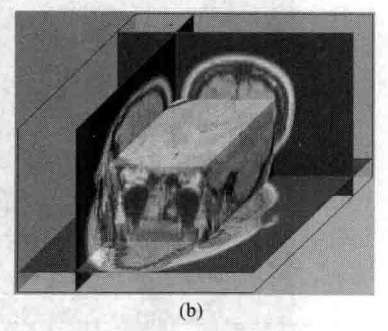

(b)

图 14.16 图 14.13 所示 MRI 头部数据的两个视图，它们是与体素阵列的轴垂直的两个剖面。体素取自轴向切片，因此平面法向分辨率要差于平面内的分辨率

遗憾的是，没有一种较好的方式来在印刷物中演示这种时基动画。有些儿童卡通图书会使用一种快速翻阅的方式来对印刷在各页上的图形生成动画，即只要翻阅的速度足够快，观看者便能在眼中形成连续运动的视觉效果。这种形式的动画需要很多页面，且实际上只适合于像卡通这样简单的图形。书籍和学术期刊的出版商很少采用这种方式。这里所能做的是，通过显示来自这样一个序列的一些静态图像，激发读者的想象力，进而尽可能提升真人动画的效果。许多在线网站以 Quicktime 电影方式显示了许多这样的动画。

在第 2 章中，图 2.35 显示了一组可用于显示这种“电影”的图像（迈布里奇的“奔腾的马”序列）。在线发布技术文献的当前趋势，也许会提供一种使用时间作为第三个轴来取代空间轴的媒 介，并通过运动来显示三维结构。在下面介绍旋转和其他时基显示方法时，会再次提及这种可能性。

在许多情形下，会使用序列图像来显示多维数据。快速访问一组平面是显示占据三个空间维度的数据集的一种方式。另一种有效的动画可显示完整三维数据集的一个视图，同时改变体素的不透明性。甚至对于占据两个空间维度的数据集，可以有效地使用两类信息之间的过渡。例如，使用二维地图作为路网、地质构造、卫星图像等多维信息的组织基础的 GIS（地理信息系统）数据集，就包含有不同类型的数据，且一次仅能显示一小部分。

如第 13 章中提及的那样，时间本身就是一个有效的第三维度，快速演替地获取一组覆盖三个空间维度的图像来研究结构或成分在一段时间内的变化时，可采用许多相同的分析和可视化工具。在气象节目中，通常会在几秒内显示完几小时或几天获得的气象卫星图像，进而演示云图的变化（见图 14.17）。其他领域也使用了同一方法。图 14.18 显示了由共焦光学显微镜以视频速率（30 帧/秒）录制的几幅图像。这样的数据集可组合到其 z 方向是时间的一个立方体中，在沿 z 方向的平面中切割这个三维体，就可研究其中的变化。

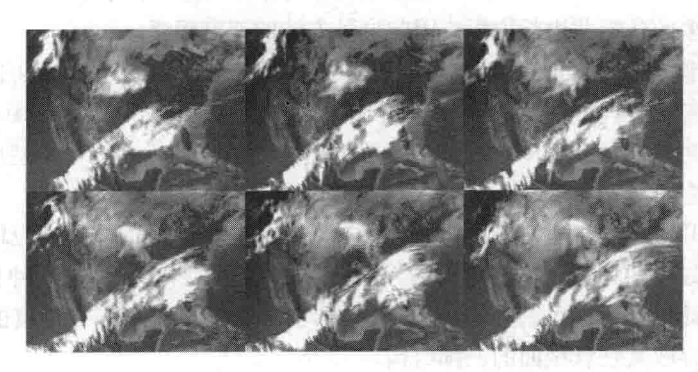


图 14.17 以 3 小时间隔拍摄的一组卫星图像，显示了美国上空的云图变化

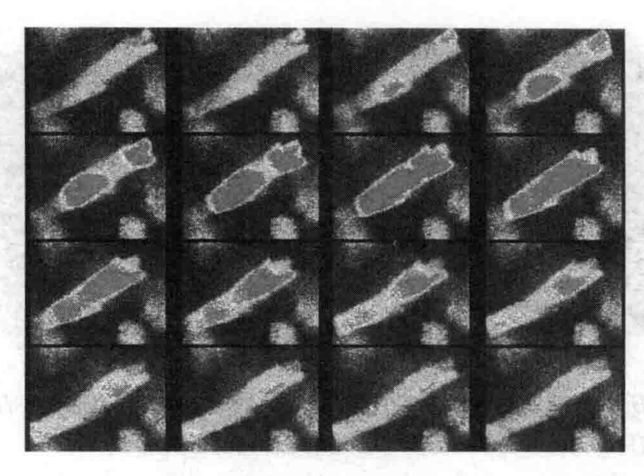


图 14.18　以钙指示剂荧光 3 加载的成年老鼠心房肌细胞的一组图像。这些图像由共焦光学显微镜以视频速率（30 帧/秒）录制

14.8　任意剖面

使剖面严格垂直于 x 轴、y 轴或 z 轴是有限制的，这样做可便于存取内存（或磁盘）中的体素数据并创建显示图像。如果选定了一些任意的平面方向，就须找到最靠近该平面的那些体素。至于第 4 章中介绍的图像扭曲、拉伸和旋转情形，这些体素通常并不恰好位于平面上，也没有规则的网格状间隔来使这些体素本身形成一幅图像。图 14.19 给出了穿过立方体素阵列的一个剖面的例子，其中揭示了部分不同的尺寸和形状。这些变化使得在平面上显示体素内容变得更为复杂。

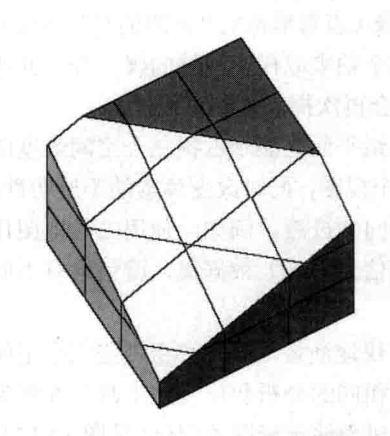


图 14.19　一个平面与立方体素阵列的交叉口。视图垂直于平面，显示出交叉区域不同的面积和形状

可用的解决方案要么是使用最靠近剖面的体素，画出它们的位置并展开它们以填充出现的任何间隙，要么在剖面中设置网格点，然后由最近邻的体素（可能多达 8 个）内插。至于旋转和拉伸的情形，这两种解决方案都有不同的缺点。使用最近邻的体素可保留亮度值（或任何体素值），但会使得边界变形并产生阶梯状图案或混叠。内插会使得边界看上去是直的和光滑的，但平均操作可平滑亮度值，使得阶梯状图案模糊。因为必须要在阵列中定位更多的体素并进行内插，因此这种方法的速度很慢。

在任意方向连续移动一个剖面来生成动画时，需要计算并存取数据（它们可能并不都在活动内存中），甚至寻址计算和值的内插，也需要计算并存储数据。如果不能实时进行这种计算，那么可存储每个平面位置的图像来进行回放。这一过程适用于展示用户已发现的内容，但不适用于研究三维数据集来发现意料之外的内容。

平面反切的最终用途是实时且动态地改变截面的位置和方向，进而即时响应场景的变化。第 2 章中曾指出，人类是通过反复考虑来研究事物的。这种反复考虑是研究物体的一种自然方式，但需要容纳所有的体素数据的计算机内存、高速处理器、显示器，以及提供所需数量自由度的用户界面。图 14.20 显示了一个可以实时存取其任意截面的三维结构。

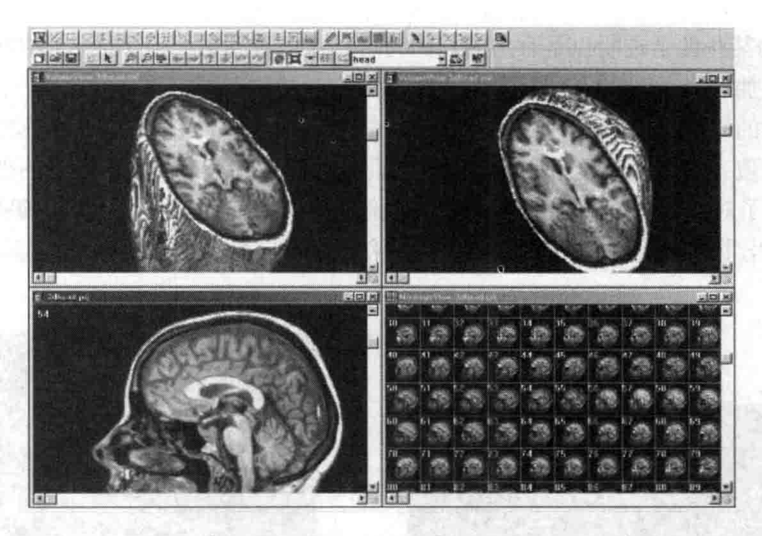


图 14.20　台式计算机上显示的剖面图像、三维重建结果以及任意剖面

定位任意一个平面的方式有两种，这两种方式会产生用户感觉不同但实际相同的结果。一种方式是相对一个固定的三维体素阵列移动平面。这是可以做到的，如沿 x 轴、y 轴和 z 轴拖动平面的角点，或先找到中心点，然后围绕通过该点的两个轴倾斜平面。另一种不同的方法是，在使得数据集在空间中旋转的同时，使剖面始终垂直于观察方向。这种方法通过让剖面靠近或远离观察者（或等效地移动数据集），可精确地获得相同的信息。

这些方法的主要不同是，在后一种情形下，图像不是透视或透视收缩观看的，因此可以进行大小比较或测量。在某些应用中，获得这样的数据可能很重要。另一方面，保持数据集固定而移动剖面可帮助用户保持结构内的方向。人与三维数据集交互的方式有多种，哪种方式最为可取并不明朗。这种交互的控制机理也很原始。使用鼠标滚动多个滑块来找到角度或位置是笨拙的，但可用的三维工具很少。在大屏幕或个人头戴式显示器上观看数据集的同时，使用仪器手套“翻阅”它，是一些虚拟现实中的有趣景象，但对多数用户而言这仍遥不可及（至少要到它们取代了当前的游戏控制台）。下面会详细探讨交互式虚拟现实。

图 14.21 显示了由穿过 MRI 图像组的三维数据集的一个任意剖面所揭示的体素。阶梯状的体素外观分散了人的注意力，因此更常显示最靠近该平面的体素的值，或在这些体素之间通过内插来得到平面上每个点的亮度值，如图 14.22 所示。

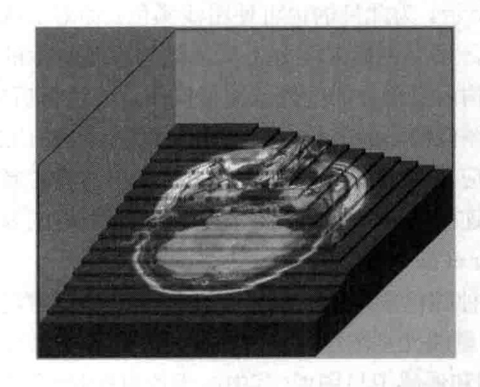


图 14.21　沿三维阵列中的各个斜面对体素采样。它表明每个完整的体素在视觉上是令人分心的，且不会产生平滑的图像

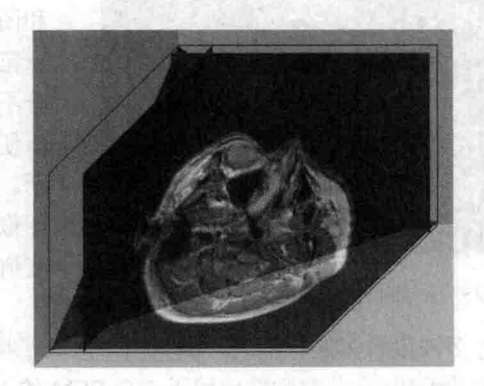


图 14.22　体素阵列中任意平面上图像像素的平滑内插

使得剖面中的某些像素透明也是有用的，因为这样就可看到第一个平面之后的其他平面，并使得数据的三维结构更加明显。图 14.23 显示了第 13 章中介绍的球形颗粒数据的这样一个例子。这组平行的切片显示为单独的平面，平面的间距在 z 方向以 2 的倍数增加。密度低于某个阈值（该阈值大致对应于颗粒密度的阈值）的那些体素会变得透明，因此就可以看到最前面剖面之后的部分颗粒。

图 14.24 显示了对人体头部的 MRI 数据集的类似处理结果。与球体相比，让哪些体素变得透明的阈值选择更为随意，因为头部内外存在许多空区域和低密度区域。但这种处理明显增强了三维结构的总体印象。

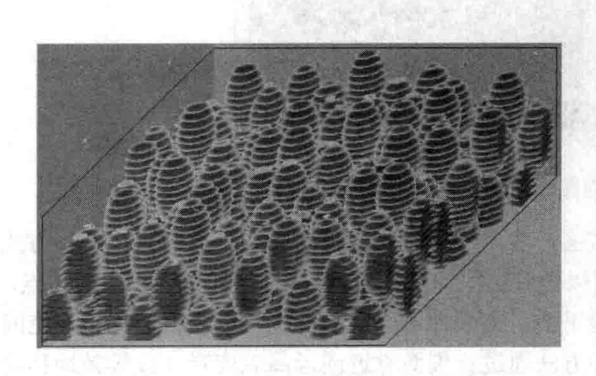


图 14.23　球形颗粒层析成像重建（见第 13 章的图 13.62）中体素层的“爆炸”视图。以透明方式显示了颗粒周围的低密度区域

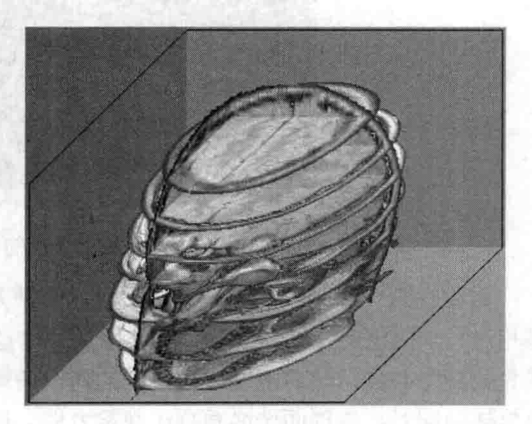


图 14.24　MRI 头部数据的几个正交切片视图，平面中的低密度区域是透明的

14.9　色彩的运用

前几章中说明了如何对灰度二维图像分配伪彩色。伪彩色有时可以区分原始图像中按亮度难以察觉到的细微变化。但如前面提到的那样，伪彩色更常会打乱图像的整体连续性和形态，进而使得图像更难解释。当然，也可对三维图像集使用伪彩色，并产生相同的结果。

颜色或阴影的一种微妙运用是，对不同的平面方向应用稍微不同的阴影，以增强三个维度上的表现力，如图 14.15 和图 14.22 所示。为此目的也可使用浅颜色。如对二维图像示范的那样，采用色标来区分不同的结构也是有用的。区分不同的结构需要单独的处理或测量操作，或进行后处理，以便对重建图像的不同区域应用不同的颜色，如图 14.25 所示。将颜色应用到三维数据集时，色标可帮助观察者了解图像之间的连续性，同时为每个物体提供某个范围内的亮度值（Parker et al.，2005）。

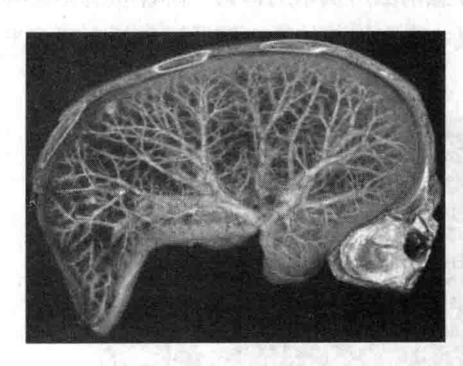


图 14.25　肺部血管的体积渲染图像，已着色

使用多种颜色来优化三维图像数据的一种最常用方法是，对多波段数据进行编码，例如由二级离子质谱仪测量的元素浓度。伪彩色的这种用途类似于二维图像中的编码，它非常频繁地用于来自 SEM 的 X 射线图和遥感卫星图像，其中的颜色可能是“真实的”，或用来表示超出了人类视觉范围的颜色，特别是红外线。处理二维多波段图像的其他工具（如计算比率）也可在三维情形下使用。

图 14.26 显示了硅晶圆中植入的几种元素的 SIMS 深度图像。叠加各幅图像并使用颜色来区分这些元素，可使得比较不同元素的空间位置非常容易。这是通过将红色、绿色和蓝色分配给三种元素来实现的，结果显示在图 14.27 中。

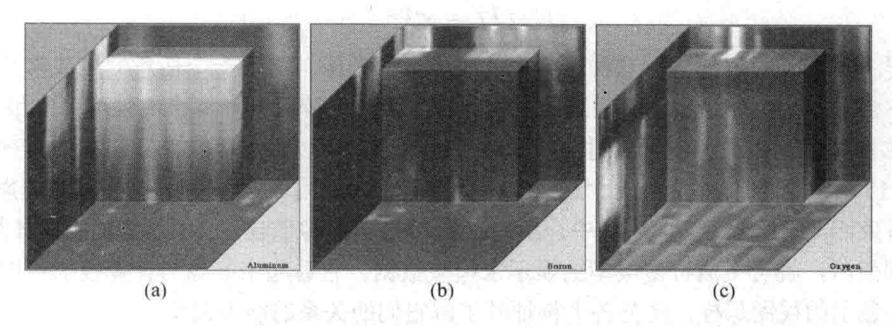


图 14.26　在硅晶圆中铝元素(a)、硼元素(b)和氧元素(c)的剖面视图，它们由二次离子质谱仪成像。图 14.27 显示了一幅彩色图像，此时所有三个元素出现在另一组正交平面上

使用不同的颜色来标注不同的结构对可视化非常有用。在图 14.28 中，使用不同颜色来显示骨架结构和血管，合并了 CT 和 MRI 数据的重建结果。合并多个数据集可为展示提供重要的补充信息（Spetsieris，1995；Ruiz-Alzola et al.，2005）。对每个数据集进行表面渲染（如下所述）和三维配准，可实现非常逼真且强大的可视化。

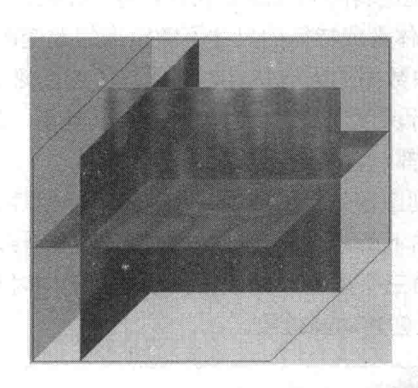


图 14.27　对图 14.26 中 SIMS 图像的元素强度着色后的效果。多波段三维数据集在 x、y 和 z 平面上切分，分别对铝、硼和氧的 256 级（8 比特）亮度分配了红色、绿色和蓝色

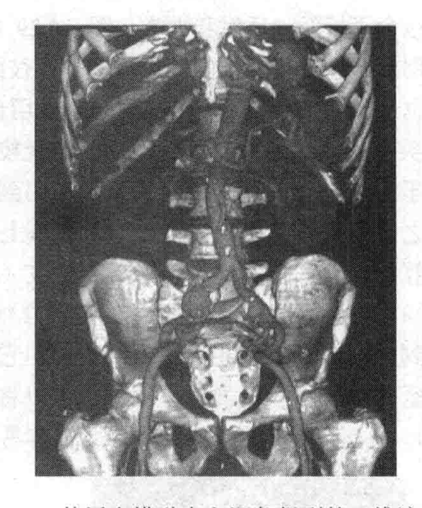


图 14.28　使用多模融合和调色得到的三维渲染图像

14.10　立体显示

即使一些区域已变得透明，对数据集切片也会遮盖许多体素阵列。能看到的只是选取的平面，且数据集中的许多信息并未用到。之前就曾指出，三维结构的拓扑性质并不会在剖面上显示。立体显示方法显示所有的三维体素信息。对于简单的结构，显示一切就是一种优势，而对于非常复杂的结构，重叠的特征和边界会使得问题非常复杂。

一种立体显示由射线追踪产生。在所用的最简模型中，体素阵列后面放置一个均匀且扩充的光源。对于从光源到显示屏幕上一个点的每条平行直线射线，沿该路径出现的每个体素的密度值，按如下吸收规则用于计算光强的下降：

$$I / I_0 = \mathrm{e}^{-\Sigma \rho} \tag{14.3}$$

对到达显示屏上每个点的射线进行这种运算，就生成一幅图像。引入一个任意的缩放常数，可将总对比度范围调整到显示范围。这种缩放很重要，因为对于大体素阵列，计算出的强度可能很小。

注意，该模型假设体素值实际上对应于密度，或对应于可以被光吸收充分建模的某种其他性质。第 13 章中图 13.61 所示图像对应于这个模型，因为它显示了使用 X 射线穿过该标本的投影视图，尽管几何是锥形射束而非平行投影。前一章中介绍的 X 射线层析成像重建，可由这样的视图计算体素阵列。得到体素阵列值后，就可生成许多类型的显示来检查数据。它看起来好像与计算投影视图相反，而且就像这些图中显示的视图那样，区分各个特征并了解它们的关系的能力很差。

这种方式的优势之一是，可以很容易地改变观看的方向。必须计算出沿每条射线出现的体素的地址。当观察方向与某个坐标轴不平行时，使用正弦/余弦值或推广的 Bresenham 直线画图算法，可有效地完成这一地址计算（Bresenham，1965）。此外，在这种情形下，计算通过每个像素的每条射线的直线段的长度，可以改进显示质量。此时的吸收规则必须在求和中包含距离 t：

$$I / I_0 = \mathrm{e}^{-\Sigma \rho t} \tag{14.4}$$

这种方法并不能处理完整的射线追踪，尽管有时它声称可以做到这一点。在一幅真实的射线跟踪图像上，包含了折射、反射和吸收。图 14.29 给出了一个例子，其中包含的阴影大大增强了三维效果。在这种应用中，计算更复杂的阴影（即特征彼此的投影）通常会占用程序的太多时间。仅使用简单的吸收方法，也可实现快速显示目的，即可以使用台式计算机交互地旋转阵列（改变观察方向）并快速显示。

这类立体显示通常是等距的而非透视校正的。换言之，体素或特征的尺寸不随距离的改变而改变。这相当于通过长焦镜头来察看场景，考虑到多数三维图像中数据的固有特性，通常这不会给观察者造成不适之感。真正的透视校正需要对屏幕上的 x、y 尺寸进行深度调整。特别是对于旋转视图，透视会增加计算量，但许多现代通用图形软件包（如 OpenGL）中都包含了透视功能。

图 14.30 显示了头虱腿关节的一幅立体视图。原始序列图像是使用透视共焦光学显微镜靠近立体体素得到的。可以看到各个肌肉纤维，但它们存在重叠。如下所述，移动堆叠图像来逼近旋转，可让观察者区分不同的肌肉群体。此外，可以将时间序列用做第三个维度，以便在移动平面的同时显示三维数据，由于不需要进行计算，因此很容易在小型计算机上创建动画效果。



图 14.29　CSLM 三维数据中分裂细胞染色体的重建。染色体是不透明的，但其周围的基质是透明的，阴影已投射到后方平面上，因此增强了三维效果

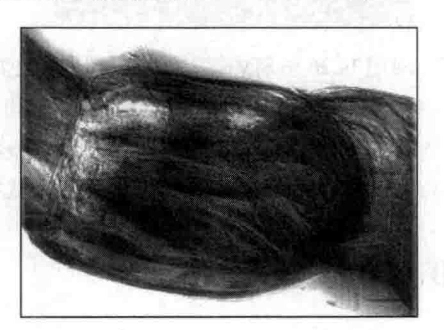
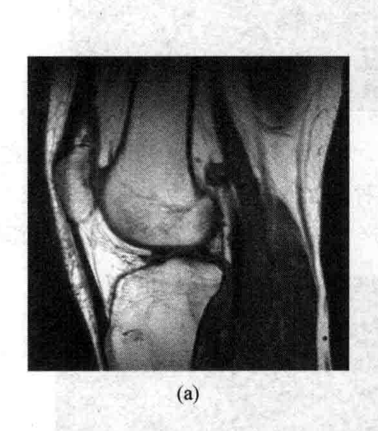
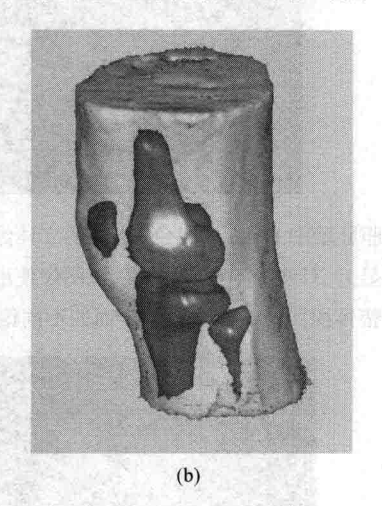


图 14.30　通过头虱腿关节的 60 幅 CSLM 堆叠图像的三维投影图像

如下所示，一种更有限的立体表示方法是，将图像的一部分显示为表面渲染模型，但将周围的组织显示为透明体。这种三维体中通常不会保留细节，它只是作为感兴趣结构的一个参考点，如图 14.31 所示。



(a)



(b)

图 14.31　观察一个表面重建：(a)通过膝关节的一个剖面；(b)多个剖面确定的骨头表面，周围的组织显示为无细节的透明体

如果寻址可简化且可忽略通过不同体素的距离变化，那么可快速地生成投影图像。图 14.32 显示了一种近似的结果，其中每个体素平面均横向平移一小段距离。这个距离可能是体素大小的整数倍，这样就可使得地址计算非常简单，但在许多情形下要求进行内插。这些平面需要与观察方向垂直，以便通过像素的所有距离相等。这类平移会给人留下小角度旋转的印象。旋转超过 30° 后，拉伸引起的三维结构变形会非常明显。但这种方法可快速地生成各个特征与深度的相对偏移，因此有助于观察者理解结构。

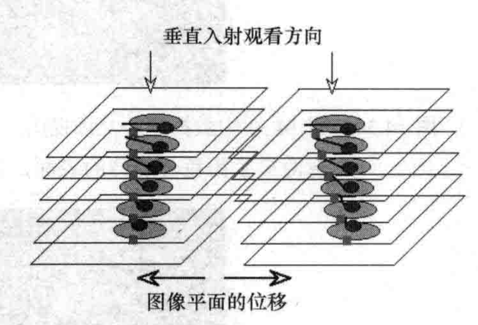


图 14.32　横向平移图像来产生旋转错觉或产生立体像对的示意图

14.11　立体观看

对于本章中的许多图像，旋转或伪旋转序列中的两幅相邻图像可视为一个立体像对。对于有些人而言，观看立体像对可能需要一种廉价的观看设备，这种设备在让眼睛保持直视的同时，可分别对两幅图像聚焦。有的人可能已掌握融合这种立体像对的技巧。但有的人可能根本看不到立体效果。例如，很多弱视人员就没有立体视觉。

对于同一标本（海胆胚胎的共焦光学显微镜图像），图 14.33 和图 14.34 显示了分别使用立体显示方法和表面渲染方法来呈现三维立体像对的例子。图 14.35 显示了在共焦光学显微镜中成像的神经细胞骨架化数据的另一个立体像对。

为左眼视图和右眼视图使用颜色通道，也可显示立体图像。例如，图 14.36 显示了仓鼠皮肤中血管的立体像对，它们是使用共焦光学显微镜在荧光模式下自动地成像的。三维重建方法使用了移动焦剖面和下面介绍的发射规则。图 14.37 中显示了使用红色和绿色合并这些图像来展示同一三维数据集的两个视图。图像是重叠的，配备有适当红色和青色滤色镜的观看者，很容易就可看到合并后图像的深度信息。

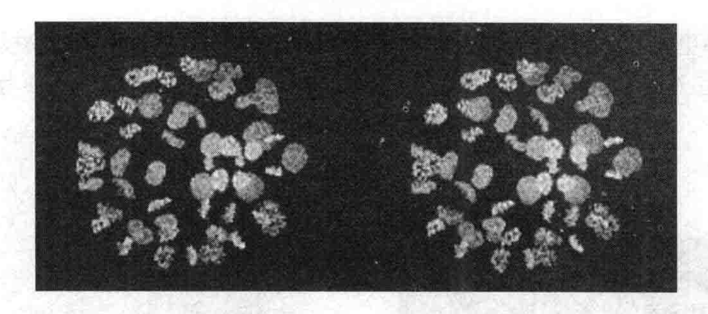


图 14.33 海胆胚胎中 Feulgen 染色 DNA 的立体像对。这些细胞形成了一个空心球体（在立体图像中很明显），有些细胞正在分裂。该图像使用一种发射模型显示了“射线投影”或“射线追踪”的一幅体积图像，各个光学剖面的不同偏移产生了立体效果。同一数据的表面图像见图 14.34

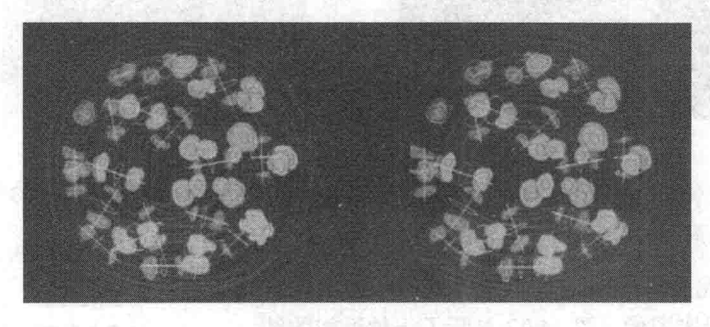


图 14.34 图 14.33 中数据集的立体视图，表面已渲染并着色。表面图像显示了每个剖面中的轮廓，这些轮廓渲染了染色体群体的表面，但掩盖了后面的任何内部细节或结构。还显示了胚胎的轮廓线

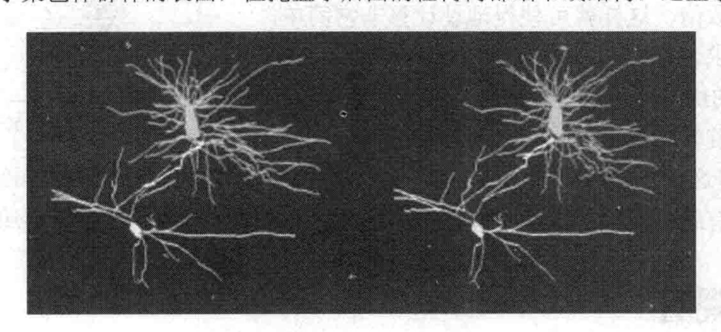


图 14.35 10 天大老鼠的海马中，两个神经元骨架线的立体图像，显示了分支和三维关系（Turner et al.，1991）

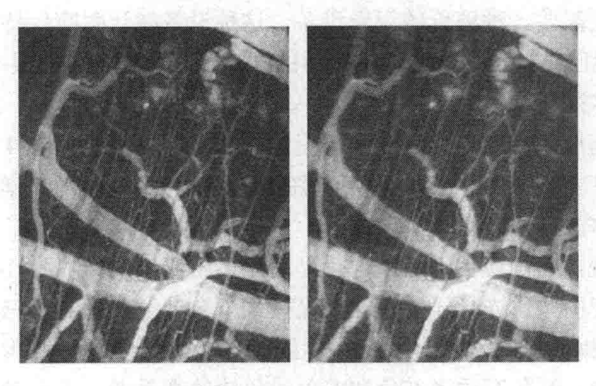


图 14.36 共焦光学显微镜的多个焦平面图像的立体视图，显示了（已注入仓鼠脉管系统并在皮肤上实时查看的）荧光染色剂发射出的光。该图像也显示在图 14.37 中

投影或射线追踪穿过一堆图像，可由多个剖面构建这些立体视图，如上所示。由共焦显微镜（其景深较浅）得到的各个剖面，可产生一组连续的体素平面。也可使用那些能产生连续薄剖面阵列的其他技术。

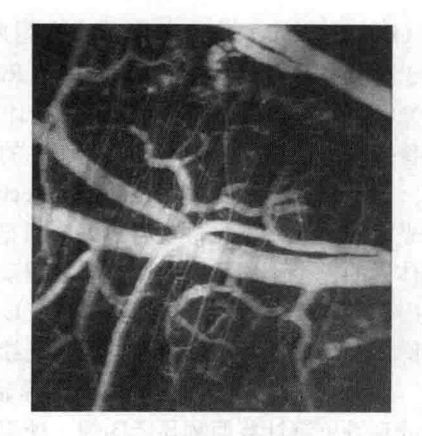


图 14.37 在图 14.36 中所示的相同图像对的立体对，对不同的眼睛观看分别用红色和青色。这允许使用眼镜（红色镜片在左眼上，绿色或蓝色的在右眼上）用垂直的眼睛朝向来观看

第 6 章中介绍了如何在频率空间处理图像来消除伪影（如模糊）的方法。三维情形下，这一过程是相同的，只是点扩散函数和傅里叶变换是三维的（Yoo et al.，2006；Hamed，2006；Tadrous，2009）。图 14.38 显示了在创建立体像对时，使用维纳逆滤波来实现这种去模糊的例子。这些图像是由 90 个光学切片重建的，切片厚度为 50μm（非常接近于光学器件的景深）。对图像逆滤波可消除多数伪影，并生成清晰的立体像对（Lin et al.，1994）。执行同类型锐化的另一种方法是应用一个迭代过程，该过程使用相邻图像来估计每个平面中的伪影。这个 Van Cittert 滤波器要求两侧的相邻平面数是点扩散函数维度的两倍，此时即光学聚焦深度（Jain，1989）。因为一次只使用几个平面，因此这种方法可能要比整个三维去卷积的速度快一些（计算机内存装不下整个数据集的情形下，会使用三维去卷积运算），即使这种方法是迭代的。

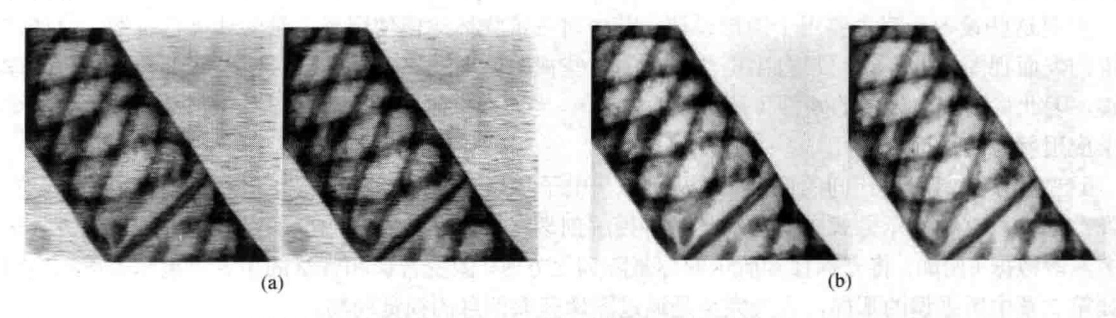


图 14.38 使用维纳滤波锐化三维图像的过程。立体像对是通过射线追踪两个投影生成的，投影的夹角为 ±2.4 度，并通过厚度为 50μm 切片中的 90 个连续光学切片。样本是水绵属植物：(a)原始图像；(b)维纳逆滤波后的图像（Lin et al.，1994）

使用配备了偏振光滤镜的两台投影仪并叠加两幅图像，可实现立体像对的投影。佩戴偏光镜的观看者可以看到立体效果，并且仍然可以使用颜色。这种显示方法需要特殊的镜面投影屏幕，这种屏幕可反射光线，但不会损失偏振，最适用于与屏幕中心正对的观看者。尽管人们有时会使用这种方法来展示三维数据，但在必须先制作摄影幻灯片时，交互式地探索三维数据集并不实用。偏振也可在单台计算机显示器上使用，详见下面的讨论。

两幅图像视角上的差异，一定程度上可通过任意控制视觉印象深度来调整。可使视角对应于人类视觉正常观看的典型会聚角。使用典型的 7.5 厘米瞳孔间距和 0.5 米的观看距离，实际的会聚角是 8.6 度。因此，深度的判断就取决于大脑对观看距离的解释，这种解释基于人眼到图像的焦距以及眼睛的会聚角度。如果会聚角是变化的，那么印象深度就可在 z 方向扩展或压缩。

14.12 专用显示硬件

某些情形下可以为三维图像分析使用专用显示硬件。全息摄影可从不同方向观看到逼真的三维显

示（Blackie et al.，1987）。计算全息图来生成这样的显示在实验上是成功的，但这种方法因太慢而不适合于交互（远逊于《星球大战》中由 R2-D2 产生的莱娅公主形象）。目前，最好的全息图是使用激光相干光和高分辨率胶片来显示的。为让计算机产生实时显示，可让平板显示器所用的多个屏幕（如 LCD）来替代胶片。但光调制（相对强度）的分辨率和控制并不完善。

另一种定制方法称为变焦镜（Fuchs et al.，1982）。在显示 CRT 上，一次画出三维阵列中体素的一个平面。用户并不直接观看屏幕，而是观看镜面反射。安装在扬声器音圈上的镜子可以移动。每当画出体素的一个不同平面时，镜子就移动一小段距离。这种运动改变了观看者眼睛到屏幕的距离，进而给出深度印象。为使绘图速度提高（以便整组平面能以 30 次/秒的速度重绘），这种技术通常只是画出轮廓，而不画出整个体素数据集。连续的轮廓可被眼睛感知为三维表面。

另外，实时查看三维计算机图形显示的最新进展是使用立体像对。计算机为方向稍微不同的数据集或观看位置计算两幅显示图像。使用 120 幅图像/秒这种高速显示硬件，交替地显示这这些图像。然后使用专用硬件来让每只眼睛仅能看到正确的图像，图像的显示速率为 60 次/秒，这种速率可以消除闪烁（人类看到连续图像而非离散图像的最低速度通常是 20 帧/秒，商业电脑的帧速率通常是 24 帧/秒，欧洲电视的速率是 25 帧/秒，美国电视的速率是 30 帧/秒）。

在显示器上安装一台液晶设备，可进行视觉切换，即快速地切换透射光的偏振方向，以便让观看者能通过含有偏振片的眼镜进行观看。第二种方法是戴上含有有源液晶装置的特殊眼镜，这种装置可在透明和不透明间快速切换。计算机的同步脉冲会使得眼镜随所示图像进行切换，进而使每只眼睛都能看到正确的视图。

所有这些设备一直主要用于图形设计，其中对三维物体建模使用了大量的计算机资源，以便产生渲染的表面视图，并允许用户自由旋转和缩放这些视图。由于许多学科均对使用三维计算机图形学感兴趣，因此肯定会出现新的硬件（及相应的软件）。当游戏机制造商发布三维显示产品时，这种技术的科学应用就为期不远了。

这些显示工具可以不同的比例尺［从纳米（电子和离子显微镜）到千米（地震勘探）］来显示三维图像数据。以这种显示方式来向医科学生讲授解剖学，效果非常不错。也可以非视觉的感知方式来处理多波段数据（例如，将光标移到所示的体素阵列上方时，改变音调和音量的声音可揭示密度和组成），但如第 2 章中所强调的那样，人类完全是通过图像获取信息的视觉动物。

对于复杂的计算机图形显示，目前并不存在最好的输入和控制设备。响应鼠标水平或垂直移动的视点旋转或平移，只能是一种粗糙的控制。可被猛击、挤压等的多点触控表面才刚开始用于三维图像。在三维空间中定位点、线和面，需要更灵活的设备或过程。在阵列中定位某个特定体素位置的方法有几种。一种方法是对两个轴使用一个 x、y 输入设备（鼠标、轨迹球等），并周期性地平移到不同的控件（它可能是屏幕上的滚动条），以调整沿第三个轴的距离。另一种方法是移动两个光标，例如，一个光标在 x-y 投影上，而另一个光标在 x-z 投影上。这些方法通常较为笨拙，因为需要在屏幕的两个区域间及交互的两个模式间来回移动。对光标进行彩色编码有助于显示深度位置。目前已出现允许用户在空间定位点（多个传感器通过三角测量法来求出位置）的三轴操纵杆和声波数字化仪，它们正用于游戏机上，但目前对于科学成像的用处不大。



图 14.39　虚拟现实示例，它包括一个头戴式立体显示器和可对三维物体进行控制的数据手套

沉浸式虚拟现实（VR）最先由 Ivan Sutherland（1965）提出。几十年前就已推出了虚拟现实的实用系统，且此领域的著述颇丰（Rheingold，1991）。实现 VR 的一种方法是同时使用头戴式显示器（HMD）与头部跟踪设备（见图 14.39）。向用户呈现的是虚拟世界的立体视图。它通过跟踪观看者头部的方向，画出与真实世界一致的虚拟图像透视

图。用户戴上眼镜后可看到立体像对，而头部的跟踪机制则根据观看者的移动和观看位置来控制所投影的内容（即视图）。

此外，任何系统通常都会有与所见内容进行交互的机械装置，如数据手套（Zimmerman et al., 1987）或其他高自由度输入，以支持虚拟世界的操控。例如，数据手套一直用于移动分子来研究酶的活性。辅以反馈力，这种方法可为研究人员提供分子最佳组合的丰富信息。尽管对资源非常苛求，但 VR 仍是一种功能强大的技术，并广泛用于娱乐产业和汽车设计中。这些系统通常一次只能由一位观看者使用，因此在完全实现“星际迷航全息甲板”这种效果之前，还有很长的路要走。

另一方面，我们要了解三维可视化的历史（Cookson，1994）始于由木材、塑料或石膏构造的物理模型。由一组剖面来建造这样一个模型是困难和费时的，且不能揭示所有的内部细节。计算机建模已从简单地勾画轮廓，发展到隐线消除、表面构建、光影和渲染，以及全套的体积或射线追踪方法。

14.13 射线追踪

上面给出的关于立体显示的几个例子中，进行了一种简单的射线追踪来对体素的密度值求和，并基于光通过三维数据集前后时所吸收的光来计算亮度。尽管该模型确实对应于某些成像情形，如透射光、电子显微镜和断层扫描成像，但在许多其他的情形下会使用不同的规则。

在横穿一个体素阵列的过程中，跟踪以一幅射线追踪图像中的一个像素结束的某个特定视线时，可用的变量包括：

（1）放在阵列后面的原始光源的亮度或颜色，这种光源要么是一个扩展光源（产生平行射线光），要么是一个点光源（产生锥形光束）。这种照明光源控制着透射光对最终图像的影响。

（2）使用第一个和/或最后一个体素中密度高于某个任意阈值的位置，来表示透明性。这些体素定义了可用反射光渲染的表面。必须为表面反射率，入射光源的位置、亮度和颜色等添加其他的规则。

（3）极大值/极小值或大梯度的位置，它们可为渲染定义任何内部表面的位置。

（4）沿路径组合体素值的规则。它可能是分数值的相乘，在体素值是线性吸收值时，可根据光子的比尔法则来对简单的吸收进行建模。在某些情形下，密度与衰减成正比，因此该规则能生成可解释的图像。还存在其他卷积规则，包括线性求和及保留极大值或极小值。尽管这些规则也对应于一些物理情形，但它们的最大价值在于能生成可描绘内部结构的图像。

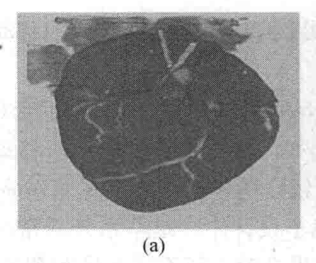
（5）体素值和源于每个体素的光的亮度（和颜色）之间的关系，这一关系表示荧光或其他发射过程。

上面第 4 条中简单提及的组合规则，对应于第 4 章和第 5 章中描述的各种图像处理工具，当时这些工具用于组合两幅或多幅图像中的像素。这些工具包括算术（相乘、相加）、排序（最小值或最大值）和布尔逻辑运算。还可以包括侧向散射，以便使得点光源在穿过体素阵列时扩散或模糊，或者进行几种模式的组合。由于测量的体素值通常与光的透射或散射相关，因此这种可能性并未得到证明。

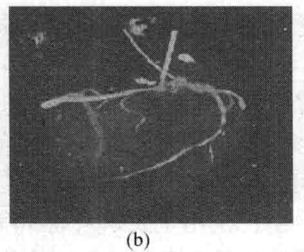
在三维可视化软件包，用户可以访问所有这些参数，并控制视图或数据集的旋转，如表面反射率和粗糙度的控制、入射光源的位置，都会影响渲染表面的外观。在沿射线路径执行透射光的处理时，体素值和吸收光间的关系是另一个控制参数。改变体素值和不透明度间的关系（线性衰减系数），或选择性地引入颜色，可显示或删除某些结构。

掌握所有这些控制和调整需要时间，而使用它们来研究数据集以找到确定无疑的新结果则需要更多的时间。

图 14.40 给出了一个示例。体素值来自 MRI 测量，并近似对应于水份含量。对不透明度、颜色和体素值使用不同的关系，可重建心脏来强调肌肉或血管。在该例中，假定每个体素既吸收放在体素阵列之后的光源所发出的光，又沿射线路径按其值和颜色的大小成比例地发出自己的光。得到的结果中，具有不同测量值的结构就以不同的颜色出现。



(a)



(b)

图 14.40　MRI 数据的立体渲染。标本是猪的心脏。改变体素值和体素不透明度显示间的任意关系，可选择性地显示心肌或血管

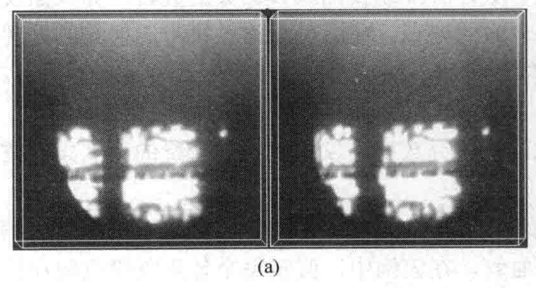
当然，若对每个体素只使用一个值，就不可能分别地对吸收和发射建模。执行双能量断层摄影来确定平均原子数和平均密度，或执行多能量断层摄影来确定每个体素中各种元素的浓度，或使用来自 MRI 的 T1、T2 弛豫时间信号，此类技术就会变得可能。它们是前几章中介绍的几个多波段和彩色成像方法的简单实现，但到目前仍不常见。在测量的一组体素值和显示的渲染间，需要采用某种任意关系来对应（由最初的成像处理测量的）主要的体素性质。

例如，在荧光光学显微镜、来自 SEM 的 X 射线图像或离子显微镜中，体素值是发射亮度的测度，它与元素浓度或化学浓度大致成正比。通过简单的射线追踪，也可立体显示这些三维数据集。替代吸收来自外部光源的光，该规则对体素值求和，来作为沿每条路径的亮度。

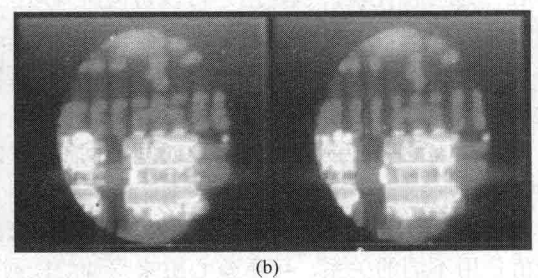
上面的图 14.36 给出了荧光共焦光学显微镜的一个应用。一种染料注入到了仓鼠的血管中并被来自显微镜的入射光激励。收集所发射的光就形成了一组具有不同焦深的二维图像，然后在堆叠中排列这些图像，就产生了一个三维数据集。此时，各平面间的间距远大于每个图像平面内的分辨率。如上所述，横向移动图像堆叠，就产生了近似的旋转和深度印象。每个体素的亮度值沿垂直列相加，就生成了每幅图像。

这个发射模型很容易计算，但不考虑其他体素沿射线路径的发射光强度的任何吸收。通常，简单的三维数据集中，每个体素只有一组数据，而没有关于密度和发射亮度的单独信息，因此不可能进行此类校正。有时，单纯根据经过的体素数量的多少来降低亮度，可逼近这种吸收效果。通常假定发射强度足够高且结构足够透明而无须这种校正，或者它不改变对结构的解释，但在任何情况下这都是定性而非定量的。

当多波段数据可用时，如图 14.26 和图 14.27 中使用的 SIMS 数据集，通过对不同的信号分配不同的颜色（最多三种），可以对发射规则使用不同的颜色。第 5 章中介绍的主成分分析可组合多个数据通道。图 14.41 显示了使用发射规则的这些数据的立体视图，它表现为一个立体像对。单色图显示了一种元素（硼），而彩色图像中则组合了多种元素。在图像中使用颜色时，需要背靠背地观看图像，而图像中密度信息［源自每个点处每个元素的 8 位（256 个灰度级）值叠加］的存在，使得我们很难融合这些图像来进行令人满意的立体观看。



(a)



(b)

图 14.41　硅晶圆中元素浓度的立体像对，它们由二次离子质谱仪成像：(a)硼元素浓度；(b)铝元素浓度（红色）、硼元素浓度（绿色）和氧元素浓度（蓝色）

对于同一数据集，使用每条视线中最前面的体素的位置（其值高于某个任意阈值），可定义内部边界的位置。如果然后使用反射光将得到的表面渲染为实心表面，就可得到另一个展示数据，如图 14.42 所示。

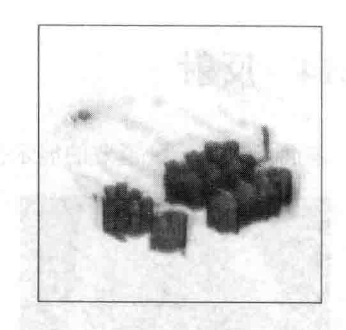


图 14.42　图 14.41(a)中同一数据的表面渲染图。富硼区域的内部表面是经阈值处理，再使用任意放置的光源渲染后确定的

显示内部表面的方法有几种。图 14.43 显示了蝙蝠螺旋耳蜗结构的几个视图（Keating，1993）。对三维体素阵列进行切片并未显示该结构是连通的。对结构内的所有体素进行阈值处理后，揭示了该螺旋结构的拓扑关系。一组剖面图像可显示为线框轮廓，也可显示为渲染的表面。旋转数据集可从不同的方向观看该结构。

显示内部表面时，可结合结构的颜色编码、立体渲染和选择性透明度（使用下面讨论的阴影）。图 14.44 显示了一个例子，即果蝇的大脑，它是使用共焦激光扫描显微镜记录的数据。三维重建表明渲染的表面在透明介质中是部分透明的，因此显示了它们之后的其他结构的表面。尽管在向观看者传递特殊信息时这是有效的，但这些显示需要预先决策和选择，而对于了解未知数据集时就不是那么有用了。

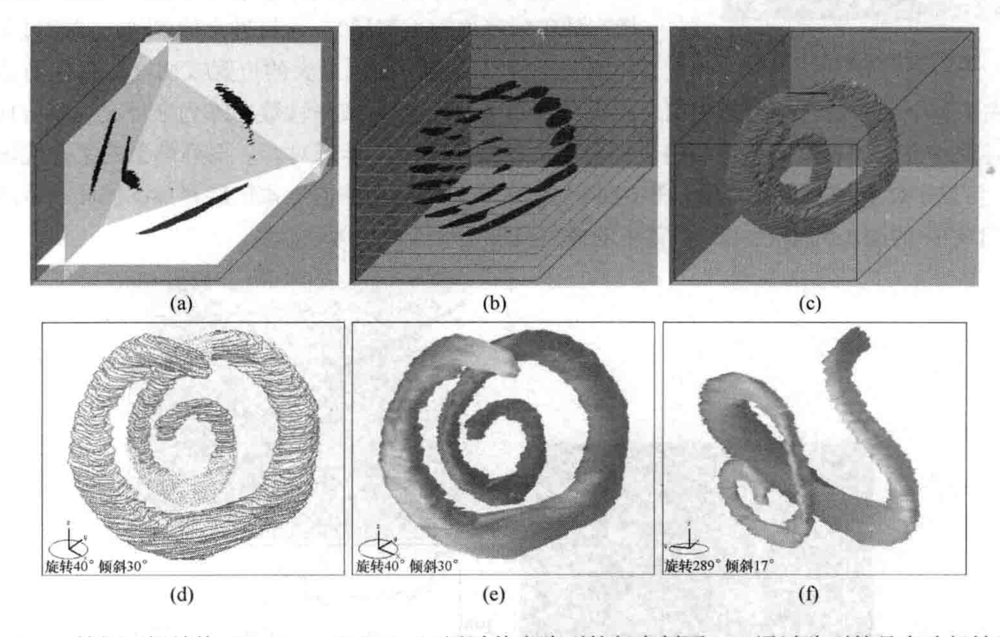


图 14.43　蝙蝠耳蜗结构（Keating，1993）：(a)通过体素阵列的任意剖面；(b)通过阵列的具有选择性透明度的平行剖面阵列；(c)体素经阈值处理后的表面；(d)描述所有体素平面中结构的轮廓；(e)由图(d)中轮子重建的表面；(f)图(e)中所示同一结构的不同视图

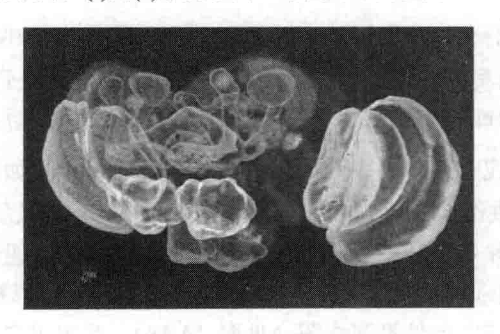


图 14.44　使用透明表面重建的果蝇大脑

14.14 反射

表面反射是一种重要的展示技术，因为人们习惯于看到并解释表面。地震波反射成图、声学显微术、超声波成像、共焦显微方法等，获取的是一个三维图像集，图像集的体素值记录来自内部位置的反射信号。图 14.45 使用超声波显示了子宫中胎儿的表面。在多数此类技术中，会生成一个体素阵列，并使用它来找到出现强反射的边界，其中的一种基质可能是不透明的。在共焦显微镜下，这种基质是透明的（空气或液体），在每个 x, y 点处的最强反射是标本表面聚焦的位置。这意味着将一组图像记录为三维数据阵列，就可定位三维空间中的表面。多数系统使用第 5 章中给出的处理方法来找到每个像素位置的最大亮度值或最大梯度，进而构造表面距离图像。



图 14.45　26 周大胎儿面部的超声波成像

产生这种显示的一种方法是，沿着数据阵列的几列数据（就像在体积射线追踪中那样），寻找最大体素值，或超过某个阈值的第一个值，或值中变化最大的位置或梯度。只保留该值或位置就生成了整个表面的一幅聚焦图像，如第 5 章中所示。旋转或平移数据阵列来更改观察方向，就可找到任何观察方向的表面点，进而将表面显示为动画序列或立体像对。对某列体素的某亮度值拟合一条曲线，可以子体素精度定位表面。图 14.46 显示了氧化铝断裂面的一幅扩展聚焦图像和一幅距离图像（来自第 1 章中的图 1.52 和图 1.53，其中的像素亮度正比于高度）。

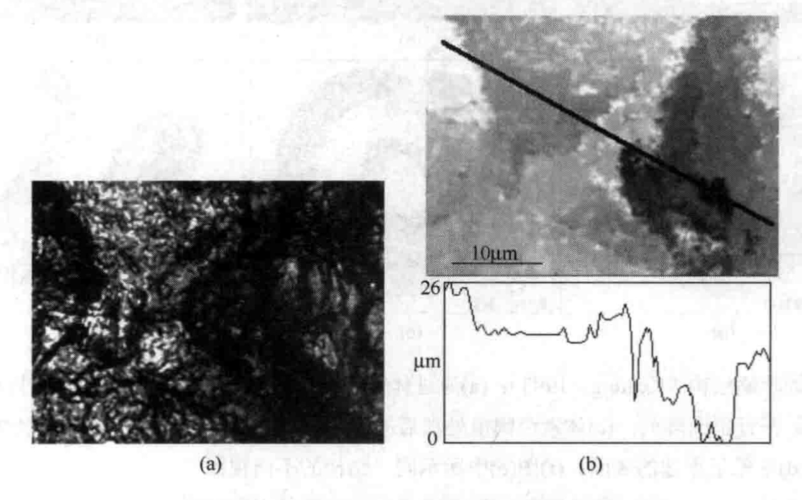


图 14.46　使用共焦扫描光学显微镜测量深度：(a)铝断裂面的扩展聚焦图像，它是通过在每个像素位置保留许多共焦平面的最亮值来得到的；(b)根据在共焦平面图像中测量的最亮（聚焦）反射，对每个像素分配一个灰度值而形成的距离图像，以及沿所示剖面的高度轮廓

距离图像的几种展示方式有助于对表面的三维形状进行可视化。如图所示，其中的一种方式是沿穿过图像的任何直线画出亮度剖面，该剖面给出了高度轮廓。图 14.47 显示了几幅其他表面高度图像（详见第 15 章）。标本是通过反射 CSLM 成像的微电子芯片，因此可由一组焦平面截面得到扩展聚焦图像和距离图像。由距离图像数据画出等高线图、网格图或渲染图，是计算机图形学中的一个简单练习。

显示表面高程的经典方法之一是等高线图（见图 14.48），其中等高线通常是以一个均匀的高度增

量画出的。这种地形图上的线条是连续和闭合的，且可在任意尺度上来使用这些方法。由于等高线图会将原始距离图像中的像素数据简化为边界表示，因此形成边界的方法与第 7 章中分割所用的方法相同。为帮助区分高程，这些线条可被标注或着色。

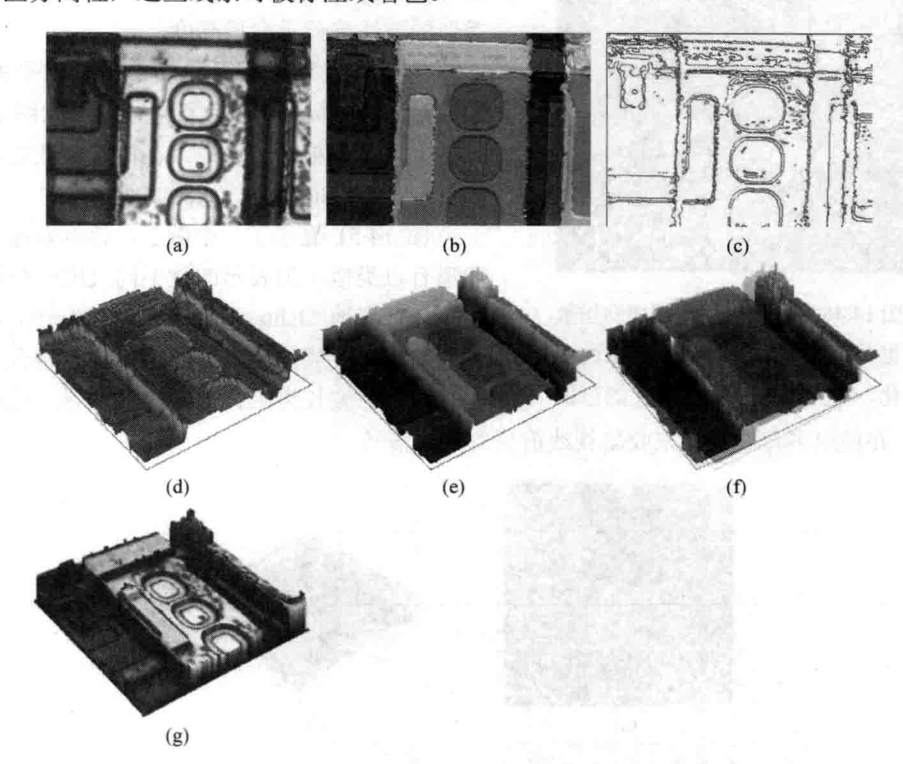


图 14.47 CSLM 表面信息的呈现方式（标本是一块微电子芯片）：(a)表面反射的聚焦图像，它是通过在每个像素处保留所有焦平面图像中的最亮值来得到的；(b)高程或距离图像，它是通过对具有最亮像素的深度进行灰度编码得到的；(c)表面高程的等值线图，已对高度值着色；(d)表面的透视校正渲染图像，添加了网格线和伪彩色；(e)与图(d)相同的图像，但已进行逼真的着色；(f)图(e)的渲染图，它是由两个视点的立体像对得到的；(g)图(d)的渲染几何，叠加了图(a)中的表面图像

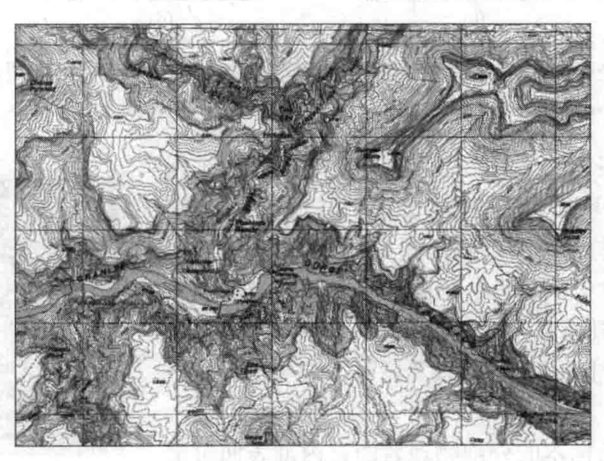


图 14.48 显示了等高线的传统地形图

将数据显示为着色的等距视图与等值线图密切相关。图 14.49 显示了图 14.46 中氧化铝断裂面的高程数据。在此图像中，使用了一个三维表示（没有透视）来将距离图像中每个像素的垂直线画为与该

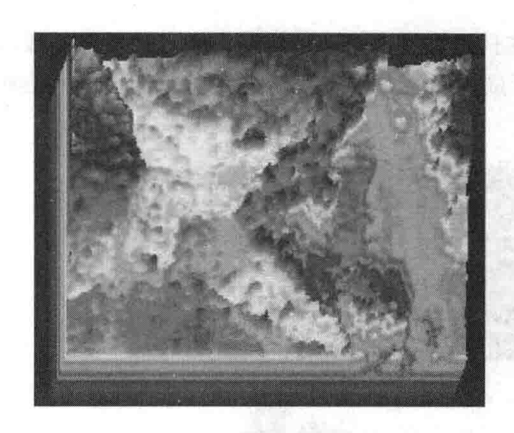


图 14.49 图 14.46 所示氧化铝断裂面高程数据的等距视图。为揭示较小的变化，使用了一个快速变化的色板

值成正比的高度。同时对图像进行了着色，以便每个点都有其灰度。使用一个伪彩色表替换灰度值，就可像等值线那样提供高程信息，并且许多地形图也使用类似的方法来显示海拔高度。

定位那些具有高于和低于阈值水平的相邻像素的像素，就可构建等值线图。但是，如图 14.50 所示，这产生了一个粗略的近似。在像素地址之间内插可产生更平滑的等值线图，如图中所示。

图 14.51 显示了一个例子，该例看起来与前一种情形有点类似，但表示的是不同尺度的数据。它是金星上伊师塔地（Ishtar Terra）的三维视图。数据来自麦哲伦号宇宙飞船的地形测绘雷达。这种合成孔径雷达向表面发射波长为 12.5 厘米的雷达波，使用回波时延来确定距离，并使用多普勒频移来收集视线前后的点的信号。

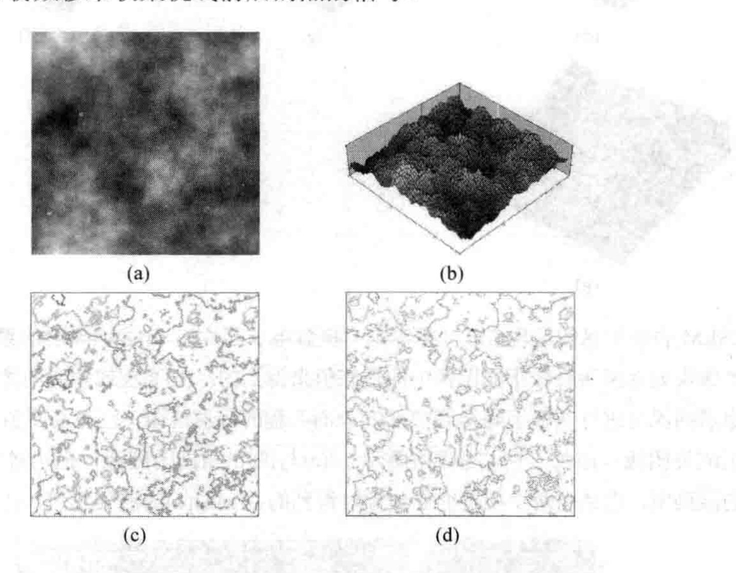


图 14.50 金属断裂表面：(a)距离图像（灰度正比于高程）；(b)着色后的格网图；(c)选取具有高于和低于每 5 个阈值的相邻像素的那些像素产生的彩色等值线图；(d)线性内插产生的等值线图

与网格或等值线地图相比，对由距离图像定义的表面进行渲染，可产生一幅更为逼真的表面图像，如图 14.52 所示。但在距离图像中，表面数据更易于定量解释。很难选取逼真的表面颜色和待应用的纹理。对于来自另一幅图像中同一区域的距离数据（如原始反射率或纹理信息），可对等距显示应用亮度值[如图 14.47(g)中所用的那样]。记录多波段图像时，这种组合相当有效。

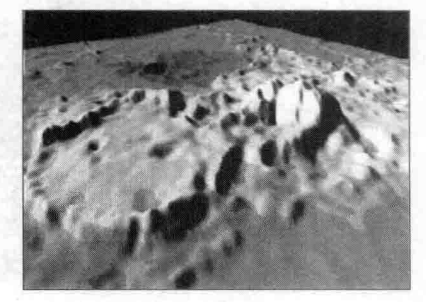


图 14.51 金星上伊师塔地的重建表面图像（Saunders，1991）

上节中讨论的表面是标本的外部物理表面。内部表面可定义为不同区域间的边界。例如，图 14.53 显示了上面所用 SIMS 例子的数据，其中在所有位置均显示了体素的深度，即硅的最大浓度。这种表面隔离为掺杂区域的形状提供了一种可视化。该图像显示为一种着色的等距显示。

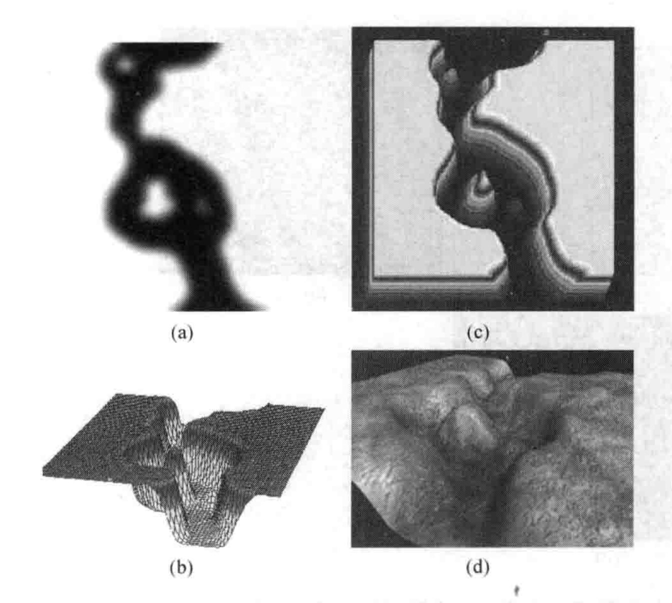


图 14.52　表面信息的显示方式：(a)距离图像；(b)网格格网；(c)等距视图；(d)渲染的地形



图 14.53　三维体内高度数据的等距显示：图 14.26 所示 SIMS 体素阵列中具有最大硅浓度的表面高度

14.15　表面

待检查的表面要么是由反射光、电子或声波揭示的物理表面，要么是样本内部的表面和获取整个三维数据集后检测的表面。使用计算机图形学来显示这些表面，类似于计算机辅助设计（CAD）中的表面显示。但典型的 CAD 对象只需要几个数字来指定它，如顶点的坐标。生成内插后的表面并计算局部方向，进而计算出许多点处的图像亮度，需要额外的计算。相比之下，此处讨论的图像数据通常形成一个完整的三维数据集，或至少是由三维数据集派生出的一幅完整的二维距离图像。因此，在显示图像的每个像素位置会有高程数据，使用这种高程数据可对图像快速进行渲染。

有些仪器可直接生成距离图像。大尺度例子包括雷达或激光测绘、立体像对计算的高度测量和声纳深度测距。在更细的尺度，测量精确加工表面的一种标准工具是干涉测量，它产生如图 14.54 所示的那些图像。亮度是高度的直接测度，而图像经适当的渲染后，理解起来更为容易。镜头伪影（图像左侧暗淡的环状结构）不是实际的高度数据，因此它们的出现令人感到奇怪。

使用上面描述的任何技术均可显示表面图像，包括线网、轮廓、等值线图和着色等距显示，它们的详细探讨请参阅第 15 章。它们理解起来都较为抽象。

一组简单的线剖面就可给出表面高程视图，且不需要进行计算，但分隔这些线会损失一些细节。因此，它有时会在诸如 SEM 或 STM 这样的仪器上用做一种直接显示方式。遗憾的是，以这种方式显示的信号可能不是真实的高度，而在 SEM 中的所谓伪地形显示可能会产生误导。在两个方向上添加网格或网线增大了有效间距，并降低了显示的横向分辨率。

逼近真实物理表面的表面图像生成，通常称为渲染法，并且需要一些计算。图 14.55 中总结了控制真实表面外观的一些物理规则。重要变量是光源的强度和位置以及观看者的位置。二者通常由到表面的法向量及光源到观察者的各个向量的夹角给出。如果表面的绝对反射率（或反照率）随波长变化，则表面会显示为彩色图像，因为有些颜色的反射率更大。

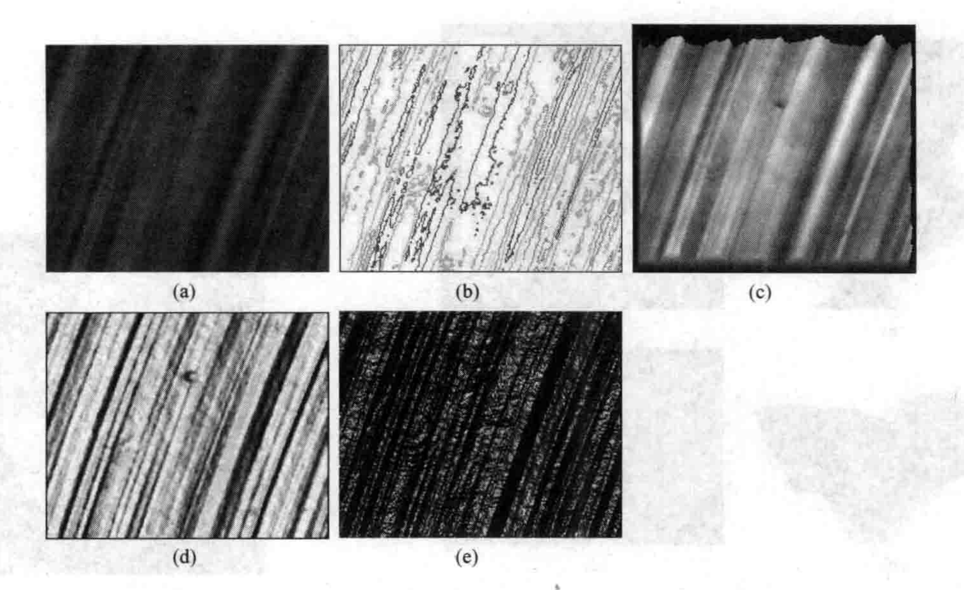


图 14.54　光学干涉仪表面高度数据的几种展示方式（标本机械加工的镍表面）：(a)原始图像，其中对高度进行了灰度亮度编码；(b)等值线图，其中的颜色编码表示了直线的高度；(c)叠加了图(a)中灰度值的等距视图；(d)渲染的表面数据，看起来像一种扩散物质；(e)渲染的表面数据，看起来像一种反射物质

$$I = k_{\text{漫反射}} I_{\text{外界}} + \frac{I_{\text{光源}}}{d}\left[k_{\text{漫反射}}(\overline{NL}) + k_{\text{镜面}}(\overline{VR})^{\eta}\right]$$

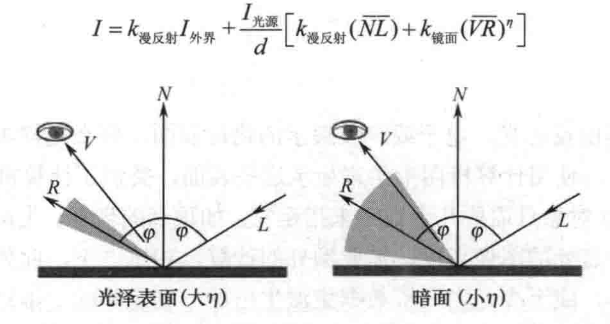


图 14.55　具有不同镜面反射特性的表面的朗伯光散射。向量是 $\boldsymbol{N}$（表面法线）、$\boldsymbol{L}$（光源）、$\boldsymbol{R}$（镜面反射）和 $\boldsymbol{V}$（观看方向）。因子 k 是漫反射和镜面反射系数。值 I 是周围和主要光源的强度，η 是描述镜面反射幅度的一个常数，主要取决于精细尺度的表面粗糙度

最后，表面的局部粗糙度控制着反射角变化的程度。角度很窄的这种扩散对应于一个镜面反射的平滑表面。较宽的角度对应于更大的漫反射。平面与电视广告中常用物体的某种较亮反射来使得物体看起来更具金属质感或更有趣。第 2 章中的图 2.47 显示了适用于科学成像的一个类似的镜面增强。

由几个点定义的典型表面（如 CAD 中的表面）可细分为几个小面（通常为三角形面），然后计算每个小面与观察者及光源的方向。再计算反射光强度，并在显示屏或其他输出设备上画出结果来建立图像。小面的数量较少时，建立图像的过程非常快，但这种图像看起来并不自然。较大的平坦小面及这些小面间的较小锐角与多数真实物体的连续表面并不相符。

对小面之间的亮度值进行着色（高洛德着色）可以消除这些陡边缘，并改善图像的外观，但需要进行内插。在各个小面中心之间内插角度而不是平均亮度值（尤其是存在镜面反射时），可实现更好的平滑。内插出的角度用于生成局部亮度值，局部亮度值随角度和位置非线性地变化。这种 Phong 着色需要更多的计算。

对于连续像素图像，可以考虑使用 3 像素组来定义三角形小面，如图 14.56 所示。由相邻像素的值（高度）的差可得到局部表面法线的角度，进而得到某个给定光源位置和表面特性的图像亮度。由于这是在像素或体数层次完成的，因此不需要对着色进行内插。

使用间隔较宽的一组剖面按这种方式完成表面渲染时，会在重建结果中产生伪影，这些伪影表现为与剖面方向平行的沟槽（见图 14.57）。使用相对较大的体素，可以生成人工块状外观的渲染结果（见图 14.58）。事先对距离图像数据应用图像处理运算，通常可改进生成的表面图像（见图 14.59）。使用计算加权平均的核进行平滑，可以生成高洛德着色。应用中值滤波器可消除表面上表现为局部尖峰或孔洞的噪声。第 4 章中介绍的滚球算子的名称来源，就在于它们对距离图像的应用。这个特定算子测试两个不同大小的相邻区域中的最小值之间的差，并消除那些差非常小的点。类比是某个半径的球体滚过表面时，不能接触的洼地。

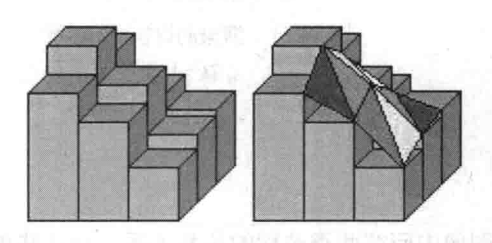


图 14.56　像素阵列离散高度值所形成表面的三角面重建说明图示

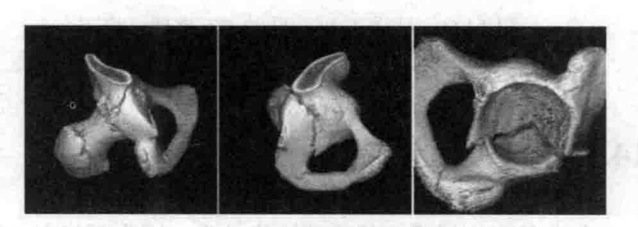


图 14.57　由连续层析成像切片图像创建的骨折骨盆渲染表面。注意由剖面间距引起的沟槽和其他表面伪影

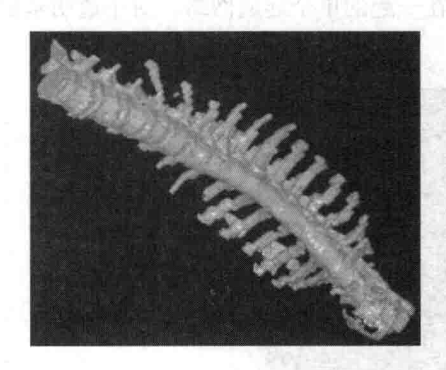


图 14.58　脊椎的渲染表面，其中粗体素尺寸产生了块状外观

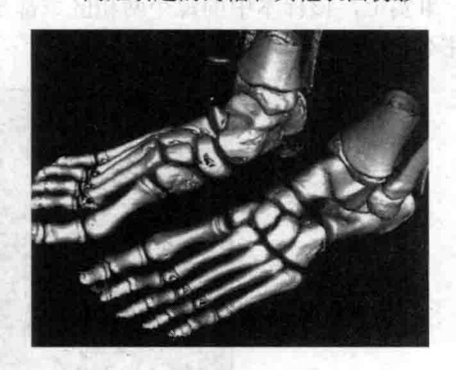


图 14.59　足部骨架的渲染图像，内插表面产生了平滑的表面外观

体素阵列内孤立结构的表面渲染可通过增加立体视图（见图 14.60）和阴影（见图 14.61）来增强。渲染后的表面图像具有真实物理对象的外观，因此便于观看者观察。但这种表面图像隐藏了原始三维图像数据集中的许多信息。需要真实射线追踪的更复杂的显示，可使得表面部分反射和部分透射，进而在显示结果中组合表面和立体信息。这种展示会在部分透明的介质中嵌入实心表面，就像果冻中的水果那样。此类展示的外观引人注目，并便于人们看到复杂的三维信息，但对于交互式地了解复杂的数据集时，会有太多的可变参数。

人们日常生活中所看到的东西是表面图像，因此表面图像很容易解释。因此，使用表面渲染来呈现立体数据是非常有用的。表面定义为某些体素性质明显发生改变的位置，它类似于第 5 章中圈定二维图像中的边缘所用的标准。获取三维数据的许多系统可以实时生成渲染的表面图像。如图 14.45 所示的图像那样，超声波扫描图像很容易被人们接受，并用于日常生活中，尽管这种技术存在固有的噪声。

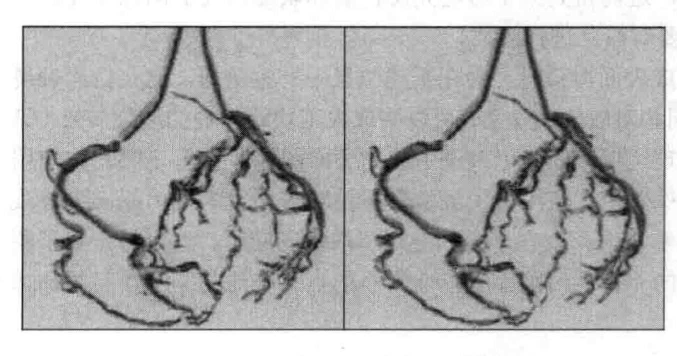


图 14.60　人类心脏动脉的 CT 重建，表面渲染并显示为一个立体像对

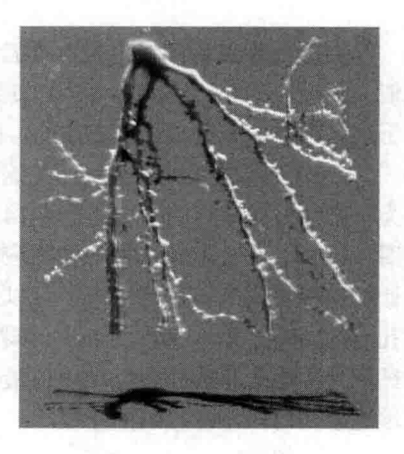


图 14.61　渲染的树状突起浮雕立体视图及其阴影

14.16　多连通表面

复杂的多连通表面最需要渲染技术，因为不能在二维图像中研究此类结构的拓扑关系。对这些更复杂的表面进行渲染是可能的。图 14.62 显示了来自离子显微镜的一个三维数据集的一组二维平面。这个两相合金样品显示了每幅图像中的许多区域。仅在完整的三维数据集中，所有区域间的连通才是明显的。事实上，该标本中的每个两相区域都是彼此紧密缠绕在一起的单个连通网络。对不透明基质切片，并不能看出这一点（见图 14.63）。

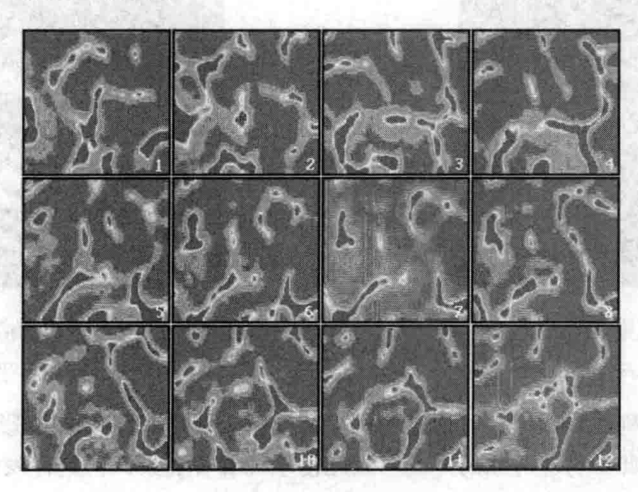
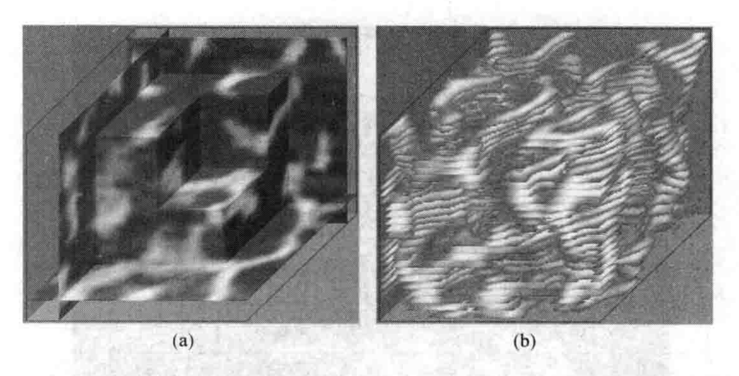


图 14.62　来自离子显微镜的连续图像，它显示了 540℃下经过 192 小时老化的 45%铁－铬合金的两相结构

该数据集的立体显示较为复杂，尤其是在因旋转而看到动画时。图 14.64 显示了使用射线追踪来生成立体显示的一些数据集的方向；按顺序快速观看，就产生了旋转的视觉效果。但这种结构的复杂性和特征位置的优先顺序限制了这种方法的用途。分开两个相间的边界，可构建一个渲染的表面，如图 14.65 所示。这种小面显示可快速画出，是支持交互式观察的一种分析工具。生成完全平滑的渲染表面（见图 14.66）需要更长的时间。

当从三维断层扫描中获得立方体素图像时，渲染内部表面会得到一种高质量的结果，如图 14.67 所示。此时，砂岩中的孔隙结构已渲染为实心物体，而基质则变得透明。这种对比度的反转对于复杂结构的可视化非常有用。



(a)　　　　　　　　　　(b)

图 14.63　图 14.62 中数据形成的三维体素阵列的剖面视图：(a)沿几个任意正交平面存储的亮度值；(b)一组平行平面上存储的值，已使得暗色体素透明

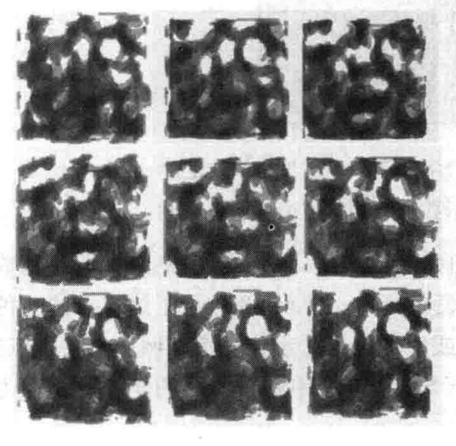


图 14.64　图 14.62 中数据的立体显示，使用了射线追踪和吸收规则。快速查看这些旋转视图就生成了可揭示该三维结构的电影

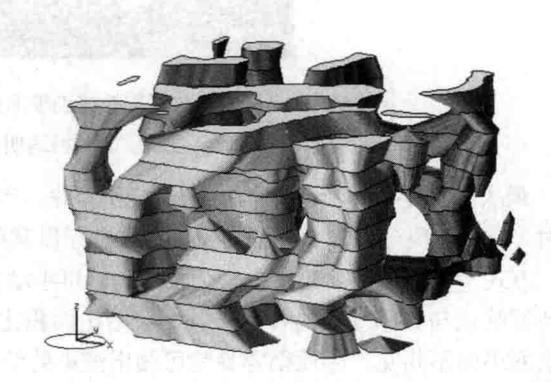
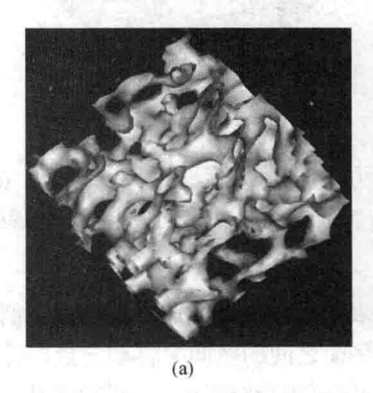
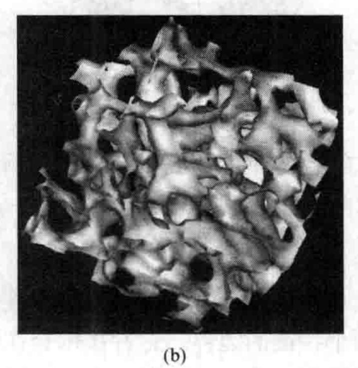


图 14.65　图 14.62 所示标本中两个相之间的边界表面的渲染，它是在平面间内插出平面小面产生的





(a)　　　　　　　　　　(b)

图 14.66　图 14.62 所示标本中两个相之间的边界表面的两个高分辨率渲染视图

表面渲染后，接着就是确定各个表面小面。最简单的一类小面是三角面。在图 14.65 所示的例子中，使用了一组较窄的矩形或梯形来连接沿每个剖面轮廓的点。由于平面间距与平面内的分辨率不同，因此需要在平面间进行内插，这就使得在深度或 z 方向表面的曲率和粗糙度不同。对于剖面形状和大小相似的那些特征，这种小面的划分相当简单。当形状在不同剖面上变化很大时，生成的小面通常不能提供表面形状的逼真视图。

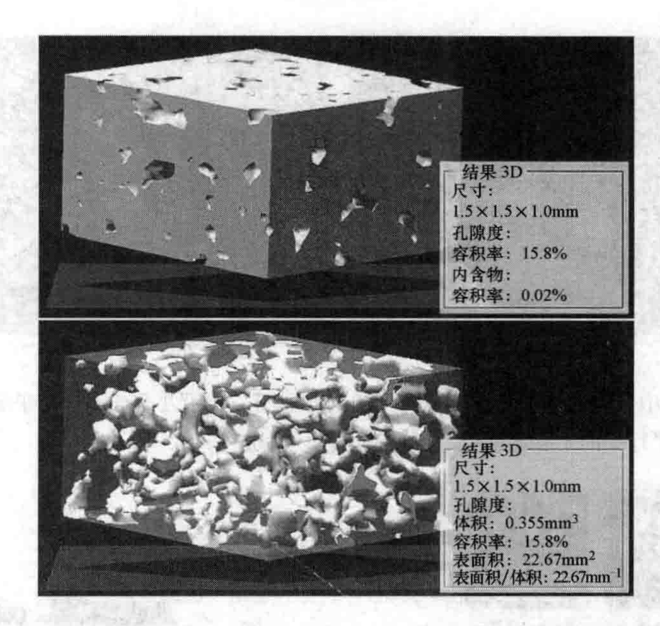


图 14.67 显微 CT 数据集的渲染表面图像。砂岩中的孔隙已被渲染为实心物体，而实心基质则变得透明，以便显示孔隙结构的连通性和复杂性

最大的困难在于处理结构中的分裂和合并。当一个剖面中的轮廓数与其下一个剖面中的轮廓数不同时，就会出现分裂和合并的情形，因此必须以某种方式将表面分开。图 14.68 显示了这样做的两种方式。方式一是，交点须在其中一个平面上。可手动定位交点或由各种算法定位交点，例如在某个点处，将特征除以与其相垂直的各个转轴，会给出面积比，它等于下一个剖面中两个特征的面积比。因为各个表面小面不相交，因此很容易就可画出渲染的结果。

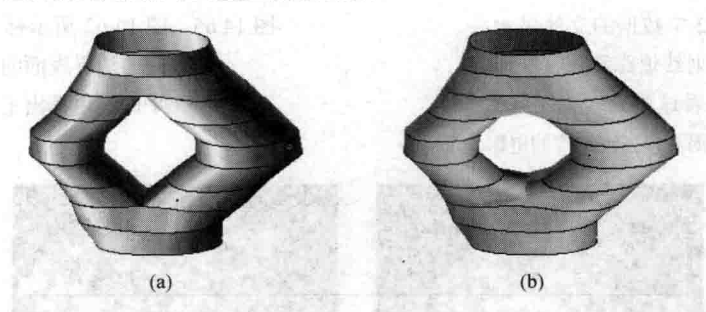
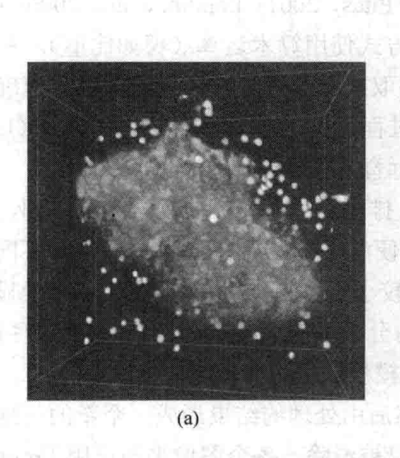


图 14.68 渲染来自连续切片轮郭的一个表面的两种方式，此时表面出现分裂或合并：(a)将一个平面分成任意多个区域，其中每个区域对应于一个分支；(b)从每个分支到下一轮廓的连续表面，它在平面之间形成了一个交点

方式二是，由一组沿特征边界出现的点来构建表面小面，特征边界是指某个剖面中的单个特征与下一剖面平面中的两个特征的边界。这种技术会在平面之间的空间中移动交点，并生成一幅相当逼真的图片，但需要的计算量很大。必须使用一个 z 缓冲区画出相交平面，z 缓冲区是一种计算机图形学技术，它记录每个图像点的深度（在观看方向），且只在它们所位于的最前面的一个表面上画出这些点。

连续切片轮廓表面渲染的主要缺点是，这些表面隐藏了它们背后的东西，即使进行旋转，也不可能看到复杂结构的所有部分。结合使用表面渲染和透明度（以便选取的特征不透明显示，而其他特征部分透明或完全透明显示，进而可看到后面的其他结构），可部分解决这一问题。图 14.69(a)给出了一个例子，其中一类特征（白色球体）是不透明的，基质是完全透明的，结构的剩余部分立体渲染为透

明的。图 14.69(b)中的例子显示了一个微电子器件的内部表面，它使用完全透明的基质渲染，因此显示了嵌入结构的几何形状。组合识别特定结构的颜色、选取的透明度和表面渲染，生成了非常有效的可视化结果，尤其是在具有旋转数据集的能力时。


(a)

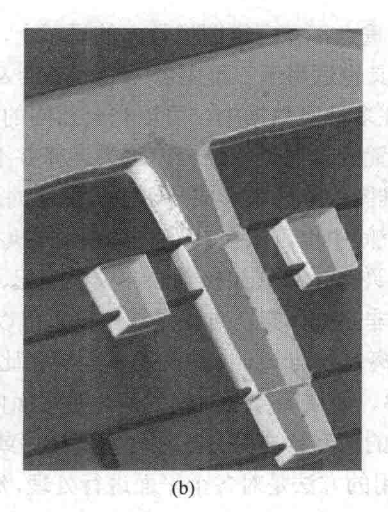

(b)

图 14.69 如文中所述，全部或部分透明的表面渲染显示了内部结构：(a)氢存储介质中的金标记颗粒（明亮）和高密度区域，它是使用 TEM 断层扫描成像的；(b)已着色为绿色的互连铜线，它包含了红色孔洞，具有蓝色蚀刻停止层，是由高角度环形暗场 STEM 断层扫描成像的

可使用多种颜色来区分复杂结构中可能出现的许多特征。图 14.70 显示了由图 14.23 中标记颜色的数据生成的约 200 个粗糙球形颗粒。使用这种伪彩色对于识别多个连通表面的连续性非常重要。

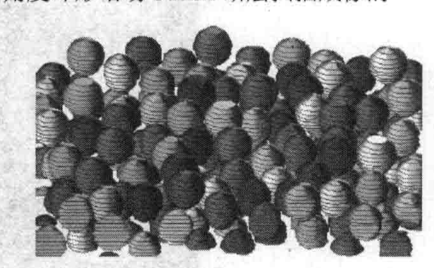
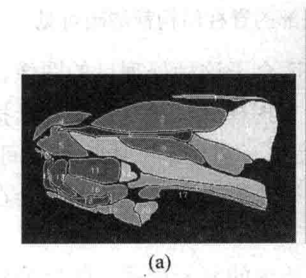
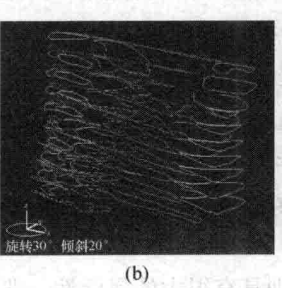


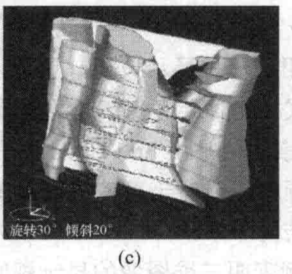
图 14.70 在图 14.23 所示的切片之间进行内插并分配颜色来区分物体时，球形颗粒的渲染表面图像

三维显示的连续切片重建并不严格限于显微学和医学应用。对纺织品和复合材料中所用纤维的研究，可使用相同的方法（Gowayed et al.，1991）。图 14.71 使用了颜色编码在日常生活的连续切片例子中识别肌肉和骨骼。访问一个当地超级市场并拍摄从牛肉一侧切下的每片烤肉，得到了这些图像。在对齐图像并对它们进行阈值处理显示出各个结构后，可渲染该切片堆叠来揭示肌肉和骨骼的三维结构。但使用第 9 章中给出的体视学过程画出网格并统计出各个切片上的点数，可更有效地求出这些结构的体积、表面积和长度。


(a)



(b)



(c)

图 14.71 牛肉一侧的连续切片：(a)已对各个特征编号的一个切片；(b)显示为轮廓的切片堆叠；(c)渲染结果，显示了具有实心表面的选取肌肉及具有虚线轮廓的几块骨头

14.17 三维空间中的图像处理

目前为止的重点都在三维图像数据的显示上，而很少提及处理。前几章中介绍的针对二维图像的多数相同处理工具也适用于三维图像（Nikolaidis & Pitas，2001；Landini et al.，2006；Ohser & Schladitz 2009）。例如，在多波段数据中，可以完全相同的方式使用算术运算（譬如比值）。每个体素值与第二幅图像中相同位置的体素值相结合。这类运算并不取决于具有立方体素的图像，但确实要求图像对齐或配准。合并二维图像阵列来形成一幅更大图像的过程，可直接扩展到三维体素阵列（Emmenlauer et al.，2009）。在三维空间中也可像在二维空间中那样执行直方图操作。

使用邻域的许多运算（如核乘法、模板匹配、排序运算等）需要立方体素阵列。在有些情形下，卷积核也适用于非立方体素，方法是调整权重值以便考虑 z 方向上到各个邻域的不同距离。这种调整仅在 z 方向的距离差异与 x, y 方向的距离差异相比较小时才起作用，例如在共焦光学显微镜中这两者之比可达 1/2 或 1/3。如果图像平面间的间距是平面内分辨率的 10 倍或以上，那么这种调整就不起作用。对立方体素形状的任何背离都会导致严重的排序或模板匹配运算问题。

此时，最实用的方法是对各个平面进行处理，然后由处理的结果形成一个新的三维图像集。图 14.72 显示了上面所用的一组 MRI 头部图像视图。头骨已被消除，各个图像平面已用 Frei-Chen 边缘算子处理，结果值用于创建立体显示。这些边缘显示了大脑的内部结构和表面皱纹。

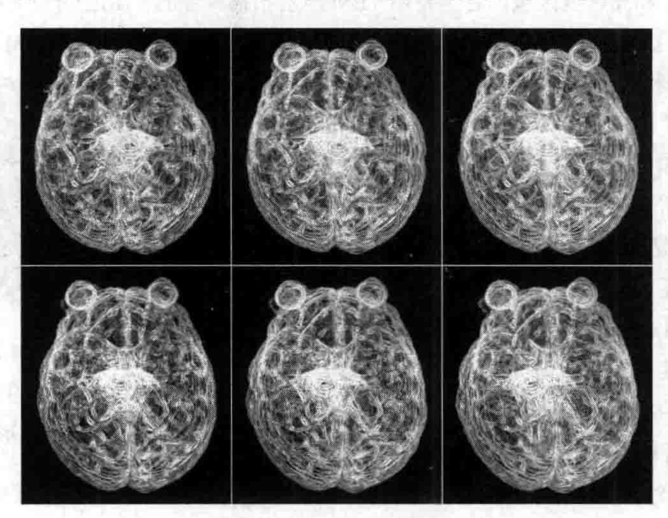


图 14.72　来自 MRI 数据集的几个大脑视图。每对图像都可立体观看，整个图像序列可形成动画。大脑顶部的裂褶到底部的脊柱结构都清晰可见

图中使用了一种显示技巧。只是不清楚哪种立体显示模式适合于这种处理后的图像。替代使用常规的吸收模式（此时透射光穿过数据阵列），对这些图像使用了发射模式（此时每个体素按其值的大小发射光）。体素的值是由 Frei-Chen 算子定义的体素边缘度。换言之，它显示了边缘在空间中的发光度。在生动的动画中，或对于可使用一对图像来立体观看三维数据集的那些读者，该图像会在大脑中形成强烈的三维结构印象。

这个示例表明三维图像的显示规则具有很大的灵活性。非传统的显示模式（特别是针对处理后的图像）对结构关系的显示通常是有效的。除了消除图像中的额外细节来显示结构外，不存在通用的准则。因此，同样鼓励采用实验的方法。

应小心进行三维数据集中图像平面的二维处理。唯一合理的是这些平面是否有首选方向，且关于

结构是随机的，或反过来说，这些平面是否有一个与结构匹配的、非常确切的已知方向。后一情形适用于涉及涂层的某些情形。在可能时，首选三维处理，尽管其计算负荷非常大。邻域的大小随立方体尺寸的增大而增大。中等大小的核（如 9×9）在二维情形下使用时可能会非常快，它对每个像素要求 81 次乘法和加法。但在三维情形下，相同的 9×9×9 核对每个体素则需要 729 次乘法和加法，而三维阵列中体素的总数要远大于二维图像中的像素数量，因此处理需要更多的时间。

对于复杂的邻域运算（如使用多个核的梯度或边缘查找运算），因为必须增加核的数量来处理更高维度的数据，因此问题会更多。例如，三维 Sobel 梯度算子使用三个方向导数之平方和的平方根。由于在三维空间中需要使用两个角度来定义一个方向，因此梯度方向的一幅图像需要两个值，且目前尚不清楚如何使用或显示它。

第 5 章中介绍的 Frei-Chen 算子（Frei & Chen，1977）是二维图像的一个非常有用的边缘检测子，添加基函数的尺寸和数量，它可扩展到三维空间中。例如，二维图像中的第一基函数（它度量一个方向上的梯度并对应于出现的一个边界）是

$$\begin{bmatrix} -1 & 0 & +1 \\ -\sqrt{2} & 0 & +\sqrt{2} \\ -1 & 0 & +1 \end{bmatrix}$$

在三维空间中，相应的 3×3×3 核是一种表面检测子：

$$\begin{bmatrix} -\sqrt{3}/3 & -\sqrt{2}/2 & -\sqrt{3}/3 \\ -\sqrt{2}/2 & -1 & -\sqrt{2}/2 \\ -\sqrt{3}/3 & -\sqrt{2}/2 & -\sqrt{3}/3 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} +\sqrt{3}/3 & +\sqrt{2}/2 & +\sqrt{3}/3 \\ +\sqrt{2}/2 & +1 & +\sqrt{2}/2 \\ +\sqrt{3}/3 & +\sqrt{2}/2 & +\sqrt{3}/3 \end{bmatrix}$$

其他核也可进行二维到三维的类似扩展。在三维空间中，还可构造一组基函数来搜索线和表面。

在许多方法中，可以使用三维处理来增强结构的可视性。在图 14.41(a)中，使用发射规则立体显示了硼浓度。但结构前面和后面部分的重叠，使得我们难以看到所有的细节。图 14.42 中的表面渲染也出现了类似的问题。图 14.73 显示了经三维处理后的相同结构。新图像中的每个体素有一个值，该值与原始图像集中 3×3×3 邻域的体素方差成正比。这些值立体显示为一幅透射图像。换言之，经过三维阵列的吸收光是边缘出现的一个测度；均匀区域看上去是透明的。这个“玻璃纸”中内部表面的可视性要远好于原始图像，且随着渲染的进行，这些表面并未遮盖它们后面的信息。

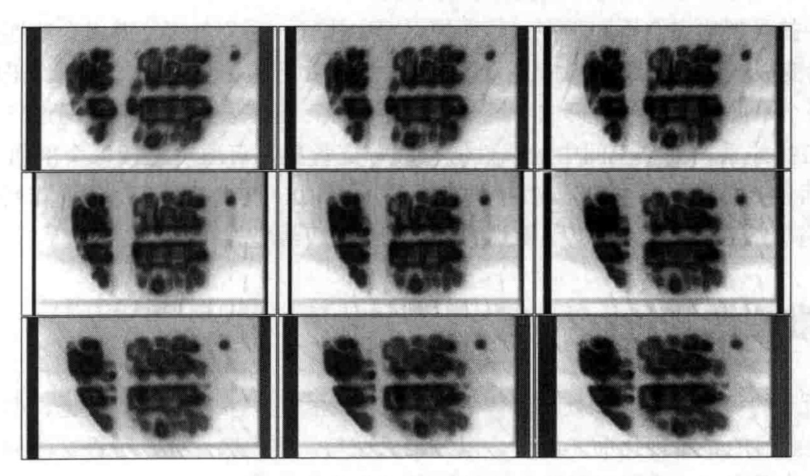


图 14.73　图 14.26 所示 SIMS 数据集中硼浓度的立体显示。这组图像通过使用像素值的局部三维方差来定位边缘，显示了伪旋转。方差的大小是沿通过体素阵列的射线的吸收光的密度值

邻域运算的计算需求比排序运算和卷积运算的计算需求更大。就像核乘法的情形那样，使用最有效的排序算法按序排列一组值的时间需求与项数并不呈正比关系，而是呈 $N\log N$ 这一比例关系。因此，较大三维邻域中的排序运算需要相对较长的时间。此外，三维空间中的混合中值滤波与第 4 章中介绍的二维混合中值滤波相比，会使用更多的子邻域来覆盖不同的方向。

对于模板匹配运算（如腐蚀、膨胀、骨架化中使用的模板匹配运算），情形会更糟。二维空间中根据邻域模式来使用查找表或命运表的有效方法，此时不再有效。在二维空间中有 8 个邻域，因此项数为 2^8=256 的一个表可以涵盖所有可能性。在三维空间中有 26 个相邻的邻域和 2^{26}=6700 万种模式。因此在确定结果时可考虑使用更少的邻域体素（例如只有 6 个面相接的邻域），或者使用不同的算法（例如能充分利用邻域模式中所有对称性的算法）。

基本的形态学运算（腐蚀、膨胀及其组合）可直接推广到三维空间中（Gratin & Meyer，1992）。使用立方体素时，必须考虑面接触、边接触和角接触邻域之间的差异。正如在二维空间中必须区分 4 连通（相接像素共边）和 8 连通（相接像素共角）那样，三维空间中的体素存在 6 连通、18 连通和 26 连通（分别共面、共边和共角）。实际应用中，许多方法会使用 6 连通这一最简单的定义。

所选像素相接标准会影响到阈值处理和特征枚举，因为不同的连通标准会连接或分开各个区域。许多基于体素的三维图像的分辨率低于二维像素分辨率的这一事实，也限制了定义特征的能力。三维阵列的阈值处理有时类似于二维切片的阈值处理，且要求这些阈值与不同切片相关联（Anderson et al.，2005；Chan et al.，2007）。在一些三维数据阵列中，基于体素值的阈值处理是一种简单而有效的方法。但在复杂的结构中，经常会使用区域生长方法，即用户先识别出一个种子像素，然后选取所有具有某个容限内的值的相连体素（通常基于 6 连通）。分割特征的另一种方法是，使用二维空间中所用活动轮廓（蛇形轮子）的一种三维扩展（气球）（Kaes et al.，1987；Cohen，1991）。

直接扩展二维欧氏距离图可构建三维欧氏距离图，三维欧氏距离图具有增强各向同性和测量表面或边界距离的优点（Borgefors，1996）。但分水岭分割在三维空间中的用处不大。

使用二维空间中的相同方法，可对三维体素阵列进行骨架化处理。也就是说，如果某个体素与背景相接或是闭（OFF）体素，就从经阈值处理后的二值结构中删除该体素，除非相接的 ON 体素确实未全部彼此相接（Borgefors et al.，1999；Halford & Preston，1984；Lobregt et al.，1980）。如果相接包括对角邻域相接、边相接和面相接，那么可以构建一个最小骨架。但是，若不能使用有 26 种可能相接邻域的一个表，则必须为每个邻域统计相接体素的数量。

三维空间中的骨架化完全不同于在二维图像平面中执行一系列骨架化并组合或连接它们。在三维空间中，骨架会变成一系列描绘结构拓扑关系的线性链接和分支。如果在二维图像平面中执行这一操作，那么连接每个平面中的骨架可形成一组薄片状的表面，这些表面会扭曲地通过三维物体，且与结构存在不同的拓扑关系。三维空间中的线性链接和分支骨架，可通过连接每个剖面中的最终腐蚀点（UEP）形成，而连接二维剖面中的分支点可构建不同的骨架。这既不同于从体素阵列的条件腐蚀直接构造骨架，也不同于从三维欧氏距离图得到的中轴变换的三维骨架。

14.18 三维图像测量

如前几章所示的那样，收集和处理图像的原因之一是由它们获得定量的数据。尽管到目前为止使用三维图像的主要目的是可视化而非测量，但三维成像与二维成像一样，可进行测量。对此类测量，可进行一些额外的观测，它们的实用性和结果的准确性非常有用。

测量大致分为两类：特定特征的测量和基于全局或场景的测量。最知名的全局测量是所选相或区

域的体积比。假设经阈值处理后选取了某个相，那么统计该相中的体素数量并除以阵列中或其他参考体积中的体素总数，就可估计出体积比。结果与体素是否为立方体素无关。统计图像平面上的像素数量可得到相同的结果，且它确实不依赖于这些平面在三维阵列中的排列方式。

第二个全局参数是所选边界单位体积中的表面积。由二维图像的测量结果求出该值存在一些体视学规则，详见第 9 章。一种方法是统计随机线（对于随机结构可使用规则网格）与边界相交的数量。另一种方法是测量二维图像中的边界长度。每个这样的值都可用于计算三维表面面积。

直接测量三维数据集中的面积看起来是一种不需要许多假设的高效方法。但在实际应用中，如何应用这种方法并不清楚。首先，边界的分辨率（特别是在边界不规则和粗糙时）严格依赖于像素或体素的大小。三维阵列可以处理的体素数量限制，会迫使各个体素大于期望的尺寸。在二维空间中，一幅 1000×1000 像素的图像需要 1MB 的存储空间，而同样的存储空间仅能容纳一个 100×100×100 的三维体素阵列。

使用更小像素来更好地定义边界并不是在二维空间中进行测量的唯一优点。三维阵列中边界面积之和必须合并计算由沿边界的每三个体素组定义的三角形面积。求和过程中必须保证能找到边界的所有部分，但沿复杂或多连通表面并不存在可以跟随的唯一路径，因此不能保证可找到边界的所有部分。对于其他全局特性，如线性特征的长度或边界的曲率，也存在类似的考虑。测量全局度量参数的无偏二维体视学工具，可有效和准确地进行测量，但并非所有情形下都会首选这些工具(Leroux et al., 2009)。

特定特征测量包括大小、形状、位置和密度测量。大小测量的例子是体积、表面积和长度（最大维数）等。在三维空间中，通过计数或二维空间中所用的类似计算，可求出这些参数。在三维空间中追踪边界也存在上述问题。但在二维图像中，特征的测量必须使用几何概率关系转换到三维尺寸。这些计算基于形状假设，且在数学上是病态的。这意味着测量或假设中的一个微小误差会在计算粒度分布的过程中放大。

简单的形状（如球体）可产生合理的结果。图 14.74 显示了图 14.70 所示球形颗粒的层析成像结果。测量二维平面切片给出了必须展开来得到球体分布的圆面积，如第 9 章所示。该结果显示了分布的一些较小不确定性，包括物理上不可能出现的尺寸统计负值（实际中被剪切为零）。但总数和平均大小与直接进行三维测量的结果相当一致，但所需要的工作量要少得多。

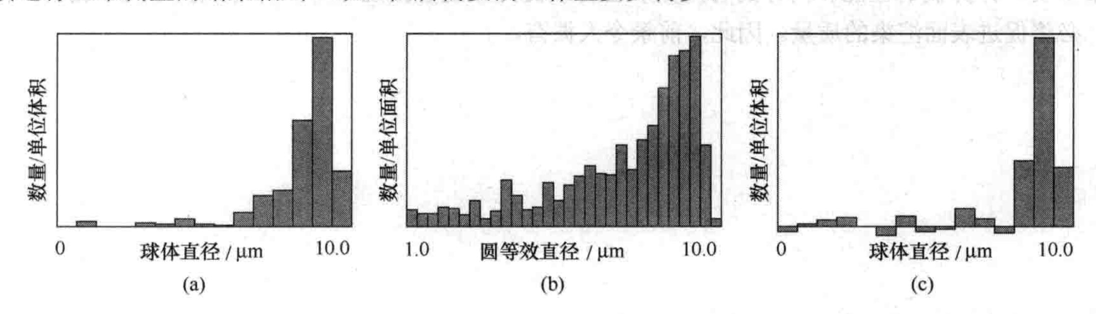


图 14.74 图 14.70 所示结构中，球形颗粒大小的二维和三维测量结果比较：(a)直接测量三维体素阵列得到的球体大小分布；(b)所有二维剖面中各个圆的大小分布；(c)展开图(a)中圆数据估计得到的球体大小分布（注意出现了负值）

当特征形状更复杂或可变时，就不能使用简单的二维方法。对于不规则形状，体积分布的均值和标准差可按第 9 章中描述的方法求出。但是，如果需要形状和大小的分布信息，则只能在三维空间中进行测量，尽管这种测量的分辨率有限。

在三维空间中确定特征的位置并不困难。在三个方向上统计像素并对矩求和，可得到质心的位置和矩轴的方向。同样，直接求和可计算出特征的密度。这些属性可以精确地求出，即使体素不是立方

体素且受体素分辨率限制的影响。

即使是在二维空间中，形状也是一个难以理解的概念。最常见的形状参数是无量纲的尺寸比，如（体积）$^{1/3}$/（表面积）$^{1/2}$ 或长度/宽度。长度可简单地定义为最长的维度，但就像二维空间中那样，正确的测量过程及宽度的定义并不是那么明显。在任何特定情形下选择哪个有意义的参数是非常自主的，要么基于研究人员的直觉，要么基于试错和统计分析。在三维空间中，由于体素的分辨率较差，因此这些值可能并不精确，但由于所用的大小参数基本是无偏的，因此准确性会更高。考虑三维形状因子时，用户的直觉很重要，因为与二维情形相比，用户更不熟悉情况。

在二维空间中确定形状的其他方法是谐波分析（这种方法展开特征边界并对结果图形执行傅里叶分析）和计算分形维数，二者均已在第 11 章中介绍过。在二维空间中可有效地求出这些参数，但在三维空间中很难求出它们。由于二维结果与三维结构立体相关（假设各个剖面各向同性的、均匀的及随机的结构样本），因此更常在各个二维图像平面上进行这些测量。

与形状紧密相关的是拓扑这一思想。拓扑是对物体或结构基本几何性质的非测度描述。拓扑性质包括环数、节点数和分支数（Aigeltinger et al.，1972）。网络结构单位体积中的连通性是一个与渗透率直接相关的物理性质。不可能由各幅二维图像直接求出三维结构的拓扑性质［尽管像 Disector 这样的最小三维探针可无偏估计出一些拓扑值（主要是数字和连通性）］。直接测量经骨架化处理而简化结构的三维数据集，可求出拓扑性质（Russ & Russ，1989）。例如，第 9 章中图 9.25 所示反手结拓扑关系的三维重建。

毫无疑问，三维成像的能力和普及性将会继续得到提升。三维成像可直接可视化和测量复杂的结构与三维关系，而使用二维成像技术无法对其进行令人满意的研究。多数生成三维图像的成像模式都很好理解，特别是层析成像重建和光学切片，但硬件和软件仍有待于进一步发展。在许多情形下，三维方法也会简化标准的制备要求。

当前的显示方法仍只能勉强地将丰富的三维图像数据集呈现给用户。新的显示算法和接口控制设备必将出现，它不仅会带动图像处理领域的发展，也会带动其他相关领域的发展，如超级计算机数据和交互式计算机游戏的可视化。人类解释复杂结构的逼真表面渲染能力，使得这些显示对于许多应用非常重要。计算机的性能和内存的容量无疑会持续提升。随着医院和工业界对此类研究设施的广泛使用，必将促进表面渲染的质量。因此，前景令人振奋。

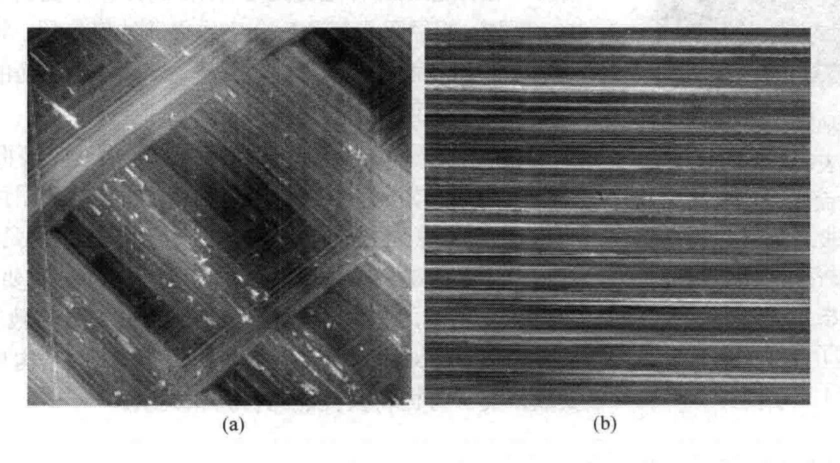
第15章 表面成像

在许多学科中，表面与块结构相比更为重要。零件间的机械作用会导致表面摩擦和磨损，许多化学反应发生在表面上，且多数现代电子器件由放在衬底表面上的薄层材料组成。物体的外观主要由其表面特性、纹理和涂层决定。在所有这些情形乃至其他情形中，科研工作者和工程技术人员都需要表征这些表面，以及制造和修改这些表面的方式。成像对获取、显示和分析这些信息非常重要。如第2章中指出的那样，人类视觉擅长解释表面图像，因此将各种类型的数据渲染为表面就成为了一种通用的可视化手段。

15.1 生成表面

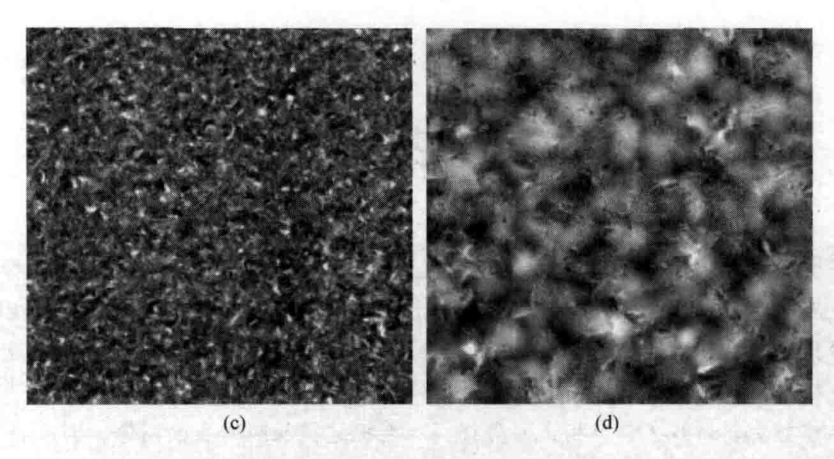
生成表面的处理有多种，有些处理严格受控，有些则相当混乱。人类生成表面的最古老方法之一是去掉一些碎料，如从一块较大的石块上去掉一些碎料来打磨出雕塑或石器。零件的现代制造过程通常包括机械加工、研磨、抛光来去掉外部材料，进而创建具有特定宏观维数和微观粗糙度的表面。

机械加工是在切削工具相对于工件移动时，从材料中去掉碎料的过程。工具刀头或边缘的形状、速度和切割深度控制着去掉碎料的方式，碎料可以是柔软的（长而连续的碎片），也可是易碎的（短而断开的碎片）。表面在某个方向上通常会出现一些较长的沟槽，其形状主要由工具的形状决定。研磨通常会导致许多切割点，同时从表面上去掉一些材料。抛光是指两个表面相对移动来去掉松散硬颗粒的过程。冲蚀（如喷砂）使用颗粒在表面上产生许多小坑。每种处理方法都会涉及塑性变形和断裂，并受其他因素的影响，如作用力、润滑剂等，它们极大地改变了所生成表面的外观和性能。当然，还有许多其他的方法，例如现代加工工艺会使用电火花放电、滚动、锻造或挤压、化学蚀刻等方法来去掉部分表面。图15.1显示了几个不同的表面。



(a) (b)

图15.1 金属表面的距离图像（每幅图像的大小均为1平方毫米，其灰度正比于高程）：(a)机械加工（快速切削）的铝表面；(b)不锈钢的研磨表面



(c) (d)

图 15.1 金属表面的距离图像（每幅图像的大小均为 1 平方毫米，其灰度正比于高程）：(c)铝的蒸汽抛光表面；(d)铜的喷砂表面。这些图像都是使用扫描探针设备得到的，设备的针尖是半径为 5 微米的钻石（续）

还有一些方法使用沉积方式来建立表面。沉积方式可以是物理的，也可以是化学的。液体（如涂料和釉料）凝固后会形成固体涂层，有时会伴随有聚合或结晶。液体在控制表面形态的模具内凝固，而某些情形下的力（如黏性力和表面张力）比实际的模具表面更重要。衬底上液体或气体的冻结，会生成非常平滑或极端粗糙的表面，具体取决于粒子和分子在表面上的运动方式（见图 15.2）。电镀通过沉积化学溶液中的原子，通常可产生光滑的表面，而颗粒的冲击沉积和聚合则会产生多孔的分形表面。之后，通过化学或物理过程，选择性地移除某些沉积层。这就是复杂多层电子芯片和微型机械设备的制造过程。



图 15.2 陶罐釉彩图像。熔融态釉质在表面上向下流动，并在表面张力下凝固为非晶状玻璃。随后，原子本身重新排列形成晶体，晶体成核并在釉质中生长，生成了有趣的图案，且改变了表面的起伏形态

关于表面，另一个需要关注的问题是其清洁度。松散分布于表面上或因静电或化学作用黏附于表面上的分子和颗粒，会破坏微电子器件中各层的沉积，因此需要净室工作环境和相应的处理方法。诸如凹坑和划痕等表面缺陷也备受关注。表面的化学改变称为污染、氧化或侵蚀，具体取决于环境，它受环境的强烈影响。某些情形下会刻意地进行这种处理，以使原始层不受其他环境因素的影响（例如，铝经阳极氧化后会生成很薄的氧化层，氧化层的化学惰性和坚硬的表面可防止对铝的进一步污染）。表面的光电特性会因极薄的污染层而出现极大的变化。

在这些情形下都需要表征表面，进而确定表面的形貌和其他特性，如化学成分或电气参数。有些表面特性数据可直接由成像方法得到。当以其他方式获取数据时，表面可视化也是一种成像技术，它依赖于人类对图像的解译来检测关于表面的重要信息。测量一般在可视化之后，它将图像数据减少到一些选取的数字，而使用这些数字可以对处理进行控制，进而一方面将表面结构与其制造过程相关联，另一方面将其性能与行为相关联。

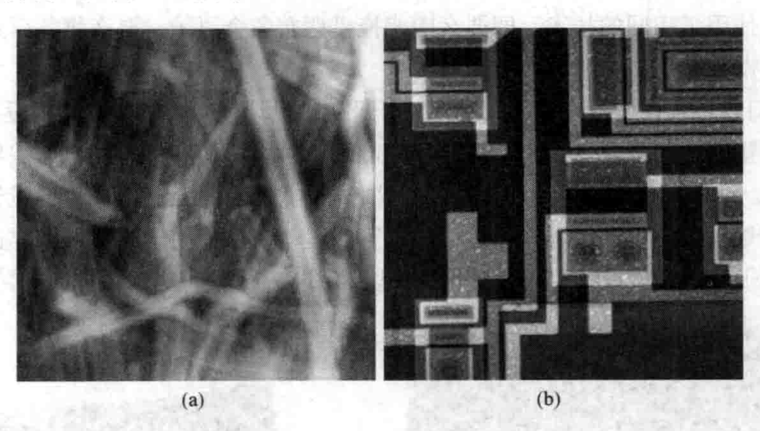
15.2 物理接触表面成像

多数表征表面的测量方法要么基于某种类型的显微镜（可提供放大的表面图像），要么基于物理接

触（使用一个较小的机械探针，或表面辐射/颗粒的散射）。这些方法可以测量成分或几何形状，包括各个薄层的厚度。研究表面时，使用了许多不同类型的显微镜和一些非传统意义上的显微镜工具（Castle & Zhdan，1997，1998；Van Helleputte，et al，1995；Russell & Batchelor，2001）。许多显微镜并不要求表面制备，对洁净度的要求也不高；有些显微镜如 SEM 则需要对电绝缘体使用导电涂层。常用方法使用可见光、电子、离子、物理接触、电子隧道、声波和其他信号来产生图像，这些图像有时与表面几何直接相关，有时则主要受表面坡度、成分、涂层厚度或微观结构的影响。最有用并且可以解释的成像方法是那些输出为“距离”值的方法，此时表面的高度直接由这些值表示，通常显示为轮廓跟踪或灰度图像。尽管这里给出的多数例子是由灰度来表示高度的，但要记住，使用相同的测量和可视化工具，也可显示化学信息或元素浓度。

涵盖某些最新设备［如原子力显微镜（AFM）和工业制造中所用的某些古老但有效的方法，如表面光度仪］的一种方法，使用了在表面上拖动的一个机械探针。放大探针的运动，逐点记录表面的高度。如果使用的是全光栅扫描，则会生成一个高度数据阵列，该阵列可显示为一幅图像，如图 15.1 中的例子所示。如果探针移动部件的质量很低，那么在相当高的扫描速率下，探针与表面（至少对于坡度为 45° 以上的表面）接触会产生几毫克或更小的力。图 15.1 中的图像每 50 秒获取一次，是覆盖 1 平方毫米面积的 500×500 个点的阵列。

工业探针设备通常使用钻石探针，针尖的半径只有几毫米，它定义了该设备的横向分辨率。使用电感、电容或干涉仪，可感知灵敏度为亚纳米级的垂向运动。由于经过了适当的校准，因此这些方法已用于定量地测量表面高度、表面抛光度和层厚等。这些设备主要使用金属和陶瓷零件，因此也可测量许多软性和脆性材料，如图 15.3 所示。



(a) (b)

图 15.3　图 15.1 中同一扫描探针设备得到的柔性或脆性表面的距离图像：(a)纸片；(b)微电子芯片

AFM 使用了一个更小的探针（Quate, 1994; Binnig et al., 1986; Gouldet al., 1990; Wickramasinghe, 1991）。扫描隧道显微镜（STM）激励一系列具有相同扫描方法和类似反馈原理的新显微镜，来得到具有纳米级横向分辨率和高度分辨率的图像。1986 年引入的原子力显微镜是检查绝缘晶体表面的一种新设备。这种设备推出时就可分辨单个原子，但直到 1993 年才出现其可分辨单个原子的确凿证据。AFM 已发展为一种灵活的设备，它为表面科学、电化学、生物学和物理学等领域的研究提供了巨大的帮助。

将硅或氮化硅蚀刻为一个尖点（将碳渲积为长而薄的尖峰），或使用碳纳米管（“巴基管”），可制造针尖为几纳米的探针（见图 15.4）。与表面光度仪相比，这种探针的横向分辨率更大。但这种针尖相当易碎和容易变形，因此为其设计了许多技术来探测表面。通常，针尖用做参考点，而表面在 z（高度）

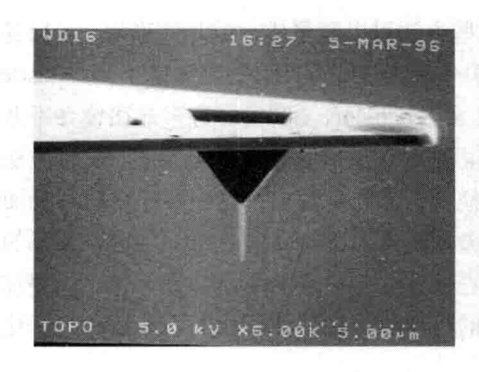


图 15.4　高分辨率原子力显微镜所用极细针尖的扫描电子显微镜图像

方向移动以与其接触。针尖与一个悬臂拱相连，悬臂拱的偏转角由背面光束的偏转角或干涉测量控制，因此很容易得到优于 1 纳米的垂直敏感度。无论是样本还是探针，均可以按 x, y 光栅图案平移，进而覆盖整个表面来创建一幅完整的图像。平移通常由压电设备实现，因此整个移动范围是受限的。

传统且仍然最常用的操作模式是使针尖与表面滑动接触。为降低探针和表面的横向力和剪切力，可快速抬起和降低探针（“轻触模式”），或转动探针求出表面材料的弹性模量进而测出横向力，或求出探针和表面间的摩擦力。在不需要物理接触的情形下，也可使用其他的模式。例如，通过范德华力，可在探针不接触表面的情形下跟踪表面。此外，有些系统使用了其他一些方法，例如在针尖接近表面时，加热针尖并测量其热量损耗，或在针尖接近表面时使针尖振动并测量特征频率的变化，或使用针尖来引导光子与表面的接触。使用振幅、相位或频移，可以测量样本表面上电力和磁力的梯度与分布，而扫描电容显微镜可测量半导体表面上少子的浓度。使分子黏附于针尖上，可在针尖于表面上移动时测量原子结合力。

最初的操作方式［其后续发展获诺贝尔奖（Gred Binnig and Heinrich Rohrer，1986）］是扫描隧道显微技术（STM），它测量半导体材料表面的电子状态。制造商和用户的实验表明，扫描探针显微镜的操作方式近乎无限，但许多技术仅适用于特殊的材料和表面类型，且所得图像的物理意义并非总是明朗的。修改表面也使用了相同的技术，即要么围绕掩模摆布各个原子，要么将各模式写入掩模，以进行微电子和微机电设备的光刻制造。

AFM 的一个问题是难以进行定量的尺寸测量。最初的设计使用开环压电陶瓷设备进行扫描，但受到了磁滞和非线性影响。任何软件校正，目前都只能修正（因不适应每个样本形貌导致的）图像变形和测量误差。因此，这极大地影响了 AFM 在密集型半导体产业中的应用。最近的设计在每个轴上都采用了一个单独的测量设备，以便闭环测量压电扫描管的运动量。使用干涉计或电容计，可精确测量一些小结构，如微电子和微机电设备、磁存储设备，以及图 15.5 所示光盘冲模机这样的结构。

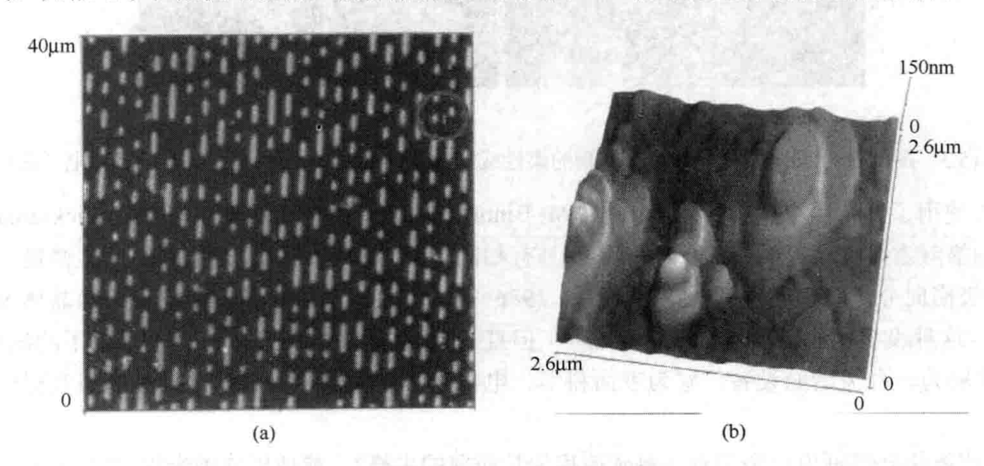


图 15.5　光盘冲模机上缺陷（圈定区域）的 AFM 图像

表面起伏的标本不受悬臂干扰，而且针尖可以访问的坡度一定时，AFM 受限于其扫描的范围和速

度。如针对表面测量方法所做的那样，具体应用要求进行某些特殊设计来突破这些限制。因此，大致了解不同技术的性能范围是有用的。图 15.6 显示了一幅图形（Stedman 图，以英国国家物理实验室的 Margaret Stedman 命名），它给出了各种表面测量技术所能容纳的横向和垂直距离范围。由 AFM 和传统探针设备检测到的最小垂直尺寸基本上是相同的，但 AFM 具有更好的横向分辨率，而探针设备则具有更大的范围。图上给出的其他技术将在本章后面介绍，但它们在处理多数表面时，均不能在垂直方向和水平方向上提供完美的距离范围和分辨率组合，也不能提供定量的准确度。

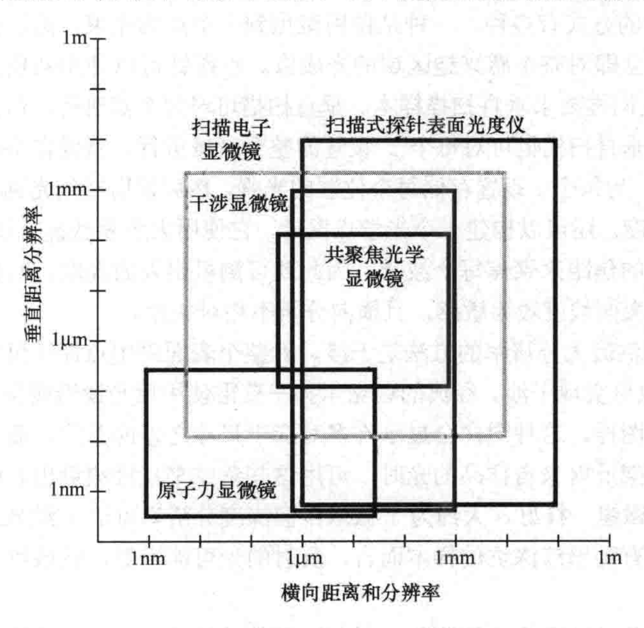


图 15.6　几种表面成像技术的典型范围和分辨率比较（Stedman 图）

15.3　非接触测量

所有探针仪器（即一次检查一个点的所有技术）的速度都是有限的，即这类仪器需要在标本上以光栅方式一次一行地移动有限数量的探针。多数探针方法还会接触表面，因此损害标本的可能性很大。AFM 一直用于创建表面形貌并对其成像，而工业探针设备通常会在软金属或聚合物表面上留下标记。对某些表面而言，最好的解决方法是制造一个可被扫描的复制品。如图 15.7 所示，塑料复制品可保留细节，并降低对原始样本的损害程度。

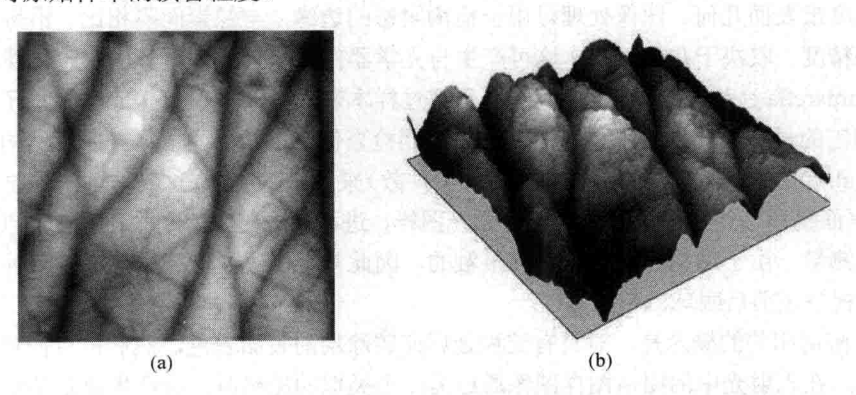


图 15.7　人类皮肤塑料复制品的扫描探针图像：(a)高度数据的灰度表示；(b)逼真的表面可视化效果

采用光作为探针可以克服速度或损害问题，但会引发其他问题。一个问题是，光不会与针尖感知的表面交互，因此测得的高度与探针方法测得的高度不一致（根据各种国际标准，后者可被人们接受）。在许多材料中，当光波穿透材料表面下方一小段距离后，会发生反射，而这种表面阻抗取决于材料的介电性质（表面污染或表面氧化层会改变这一性质）。小尺度结构也会产生斑点和干涉效应，因此会改变反射光的路径，进而生成许多错误的表面结构。此外，污染、氧化膜或几度以上的局部表面倾角，都会降低散射回检测器的光量，进而导致根本未对表面上的某些点进行测量。

使用光来探测表面的方式有多种。一种是使用聚焦到一个点的光束，然后按探针方式扫描整个样本，同时使用其他设备立即对整个感兴趣区域的光成像。点探针可以使用共焦光学来检测到样本的距离（也称焦点检测），这同样要求垂直扫描样本。垂直扫描可对每个点进行，但由于需要移动一个有限群体，因此速度很慢；垂直扫描也可对每个 z 设置的整个区域进行，就像在多数共焦光学显微镜中所做的那样。在该方法中，为每个 z 设置存储每个位置的光强，然后使用峰值光强（z 位置间内插的光强）来求出该位置的表面高度。还可以构建一个光学点探针，它使用大色差透镜来检测最强反射光的波长。因为这些透镜会以不同的焦距来聚焦每个波长，因此就可测量出表面高度。其他光学技术（如三角测量法和结构光）对局部表面坡度较为敏感，且横向分辨率相对较差。

能在最大范围内提供最大分辨率的方法是干涉，对整个表面采用点探针可实现干涉，采用单色光（通常为激光）或白光也可实现干涉。经典的迈克尔孙－莫雷使用反光镜沿两条路径发送单色光，并重新合并它们来产生干涉图样，这种图样会显示许多对应于尺寸之差的条纹，而尺寸之差远小于光的波长。当干涉仪的一个支架反射来自样品的光时，可用这些条纹来直接测量出表面高度。横向分辨率仅与收集反射光的光学显微镜一样好，大约为 1 微米，但深度分辨率可达 1 纳米或更高。但对于那些不能很好地反射光线或具有陡壁或深坑的样本而言，反射的光可能不够，或条纹的间隔可能太近，因而无法提供正确的结果。

使用单色光时，可将点的高度内插到约 1 纳米，即光波长的 1/1000。由于干涉图样在光波长的倍数处重复，为跟踪高度的变化，须使条纹间距足够大（即表面坡度须是平缓的）。许多现代系统替代使用一个波长以上的光或白光。这仅在一个焦深（路径长度相等且所有波长同样的位置）处产生干涉。对表面成像时，这意味着只有沿等高线的点是亮的。改变标本和光学器件之间的距离，可进行 z 向扫描，进而求出图像上所有点的高度。与简单的单色干涉图样相比，这需要更长的时间，且高度值的精度受限于扫描硬件的性能，但优点是可处理起伏更大、坡度更陡的表面。

间接干涉技术（如使用栅格投影来产生莫尔条纹）也可产生二维高度数据阵列。例如，考虑透过许多裂缝（如百叶窗）照射到卧室地板上的光。若地板是平坦的，则从上看时光线是直的。若地板不规则，则光影会有偏差。将这一场景缩小到几微米，就是用于查看条纹的光学分辨率，而使用相同的结构光方法可度量表面几何。图像处理可用于检测阴影的边缘，与像素间隔相比，沿每条扫描线内插可得到更好的精度。取决于几何形状，这可产生与光学器件横向分辨率近似的垂直测量精度，通常为 1 微米（Sciammarella et al.，2005），但仅适用于穿过样本表面的某些位置，除非扫描了线图样。

工具修理工的显微镜和木材平整表面的质量控制检查使用了相同的基本方法。结构光测量的一种改进（Ghita et al., 2005）通过测量投影图样的散焦（扩散）来求出深度。在这种技术的当前实现中（Masi, 2005），使用平面镜和棱镜来偏转穿过工件的激光图样，进而生成图像。这种方法的优点是可按不同方位和方向进行测量。由于结构光是快速且非接触的，因此常用于医学应用中，如整形外科中脊柱弯曲度的测量和近视矫正前后眼球曲率的测量。

与结构光密切相关的概念是，对具有淀积金属或碳涂层的表面着色，然后按特征和表面上的不规则性测量投影。在入射光中的栅格图样图像通过另一个类似的栅格时，会产生莫尔条纹，使用这种条纹中的暗线即可揭示物体的形状。这种技术对于揭示表面与理想几何形状的偏差非常有用。

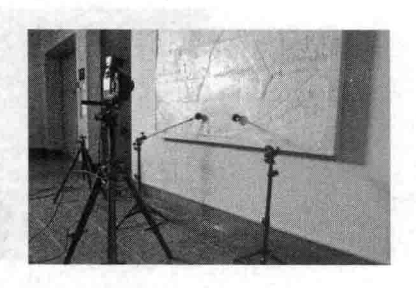
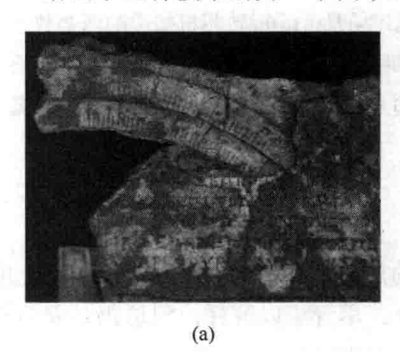


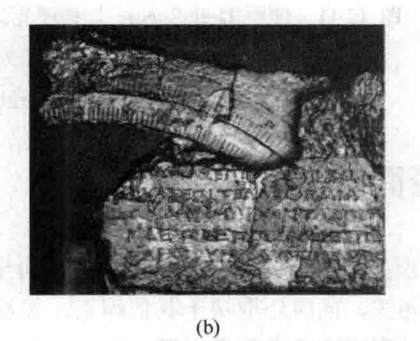
图 15.8 使用参考球（例中显示了两个）对一位埃及人浮雕表面成像，以记录表面阴影成形计算所用的光照

阴影成形（也称立体测光法）使用同一位置的不同光源拍摄的多幅图像来求表面上每个点处的局部坡度，进而计算出表面起伏。这种计算的尺度可以从微米（电子显微镜尺度）到千米（航拍尺度），但更常用于宏观物体。图 15.8 给出了一个例子，例中这种技术做了改进，即在场景中使用了一个参考球体。球体的光照给出了计算每幅图像中光照方向的信息。闪光灯与表面成不同的角度时，拍摄了 30～40 幅图像。软件计算出场景中每个位置的坡度信息，由此创建表面的多项式纹理图（PTM）（Mudge & Malzbender，2006）。这种技术特别适合记录独特的对象，如艺术和考古珍品、法庭证据以及其他类似的物体和表面（Padfield et al.，2005）。

《科学美国人》（Freeth，2009）在一篇关于安提基特拉装置（即希腊人于公元前 80 年左右发明的一种古代天文计算机）的文章中，介绍了使用这种方法的一个例子。该装置由采集海绵的潜水员于 1900 年发现，今天的学者仍不能详细了解该机械的功能，但正使用多个碎片图像复制该装置。一旦复制成功，就可显示 PTM，进而显示该物体的原始表面，或者可以进行镜面反射增强，以交互地调整光向来揭示表面细节。图 15.9 显示了镜面打光时安提基特拉装置的一部分。图 15.10 显示了 15 世纪的泥金写本，并附有参考球。通过镜面反射增强显示 PTM 后（如图像下半部所示），增强了纸和颜料的表面纹理（第 2 章中的图 2.47 给出了这种技术的另一个例子）。



(a)



(b)

图 15.9 安提基特拉装置的一部分：(a)照片；(b)镜面反射增强后，渲染的表面

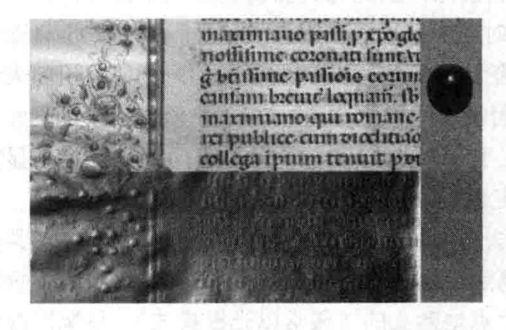


图 15.10 使用一个参考球拍摄的 15 世纪泥金写本，由参考球可导出光向并计算出多项式纹理图。图像底部显示了镜面增强后表面打光的结果，揭示了纸的纹理和各种所用颜料的厚度

在更大的尺度上，为犯罪现场存档或记录考古现场中器物位置的目的，可使用扫描激光测距仪来对整个房间成图。扫描仪在三脚架上旋转，记录到许多点的距离。同时，数码相机拍摄照片，结合这些照片与距离值，就可构造整个场景的一个模型。图 15.11 显示两幅数字图像，房间中来自墙面、地面和内部物体的许多点云，以及场景的三维重建结果，由三维重建结果可进行各种测量。整个模型可以从任何方位（包括内部点）交互式地观看，并可连续漫游于与其相连的各个空间模型（即房子的其他房间）。

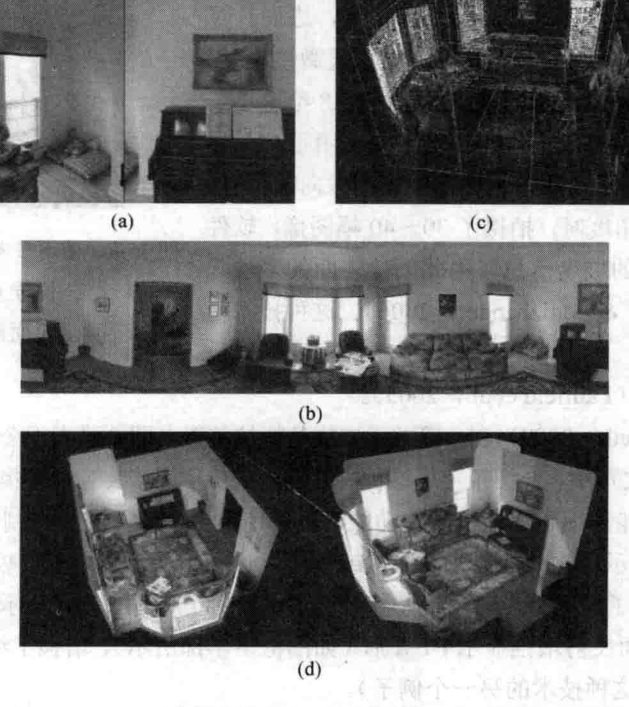


图 15.11　使用 DeltaSphere 扫描激光测距仪重建场景：(a)旋转相机拍摄的两幅数字图像；(b)合并场景可见光外观与测距仪距离数据后的完整数据集；(c)场景的三维模型，显示了激光所测量的点；(d)各个三维旋转视图及示例测量

15.4　表面的显微镜检查

多数类型显微镜所生成的图像，其亮度均与表面的反射光（或其他信号）相关。亮度仅与表面几何形状间接相关，同时还取决于其他因素，如局部成分。尽管难以解释，但这仍是表面检查的常用过程，因为该过程速度快且使用方便。

具有适当放大率的标准光学显微镜具有相对较浅的景深，因此为表面检查带来了许多问题。如果表面不平且与光轴垂直（如金相抛光试样），则不能同时聚焦。只能使用放大倍数较低的光学显微镜来检查粗糙的表面（如金属断面的显微镜观察），且无法给出关于表面几何的详细信息。使用粗糙表面散射的光图样可确定粗糙度。具有分形几何的表面会散射漫射光，进而生成一种分形图样，且表面分形维数和图像分形维数之间存在某种关系（Russ，1994；Pentland，1984）。此类表面的扫描电子显微镜图像也存在这种关系。但测量表面的整体粗糙度与确定表面上实际点的坐标并不相同。

另一方面，常规光学显微镜的景深太大，因此无法测量粗糙表面上垂直方向的重要尺寸。共焦光学显微镜的景深更浅（更重要的是，它会拒绝来自远离焦平面和焦点的杂散光），因此可生成不规则表面的真实距离图像。在共焦显微镜中，图像是一次一个点地建立的（通常以光栅模式）。每幅图像仅对应于那些具有特定焦深的点，但以许多焦深重复该操作则会生成扩展聚焦图像（此时对整个表面成像）和距离图像（此时记录每个点的高度）。垂向和横向分辨率与扫描探针显微镜或干涉仪相比要差很多，但仍适用于许多表面测量应用，如某些微电子器件的测量。对于电介质材料，使用不同方向的偏振光获取图像，可计算出表面的几何形状（Miché et al.，2005）。

由于 SEM 具有很大的景深和超高的分辨率（一般小于 10 纳米，远好于光学显微镜），因此它是检查粗糙表面的首选工具之一。此外，由于观察者非常熟悉这种设备记录的二次电子图像，因此观察者认为可以解译图像来得到几何信息。图 15.12 显示了烧结碳化钨颗粒表面的一幅 SEM 图像。这是一个相对简单的表面，它由成分均匀的颗粒组成，具有相对平坦的三角面。但高度/坡度和局部像素亮度间不存在独特或简单的关系。对于没有清晰边缘且相对平滑的表面，阴影成形方法可将亮度变化转换为坡度变化，进而提取几何信息。小尺度粗糙度、边缘、表面污染或成分变化等影响，通常会妨碍其成为一种通用方法。反向散射电子成像对许多此类影响不敏感，因此适用于某些计量应用（Paluszynski & Slowko，2009），但与当前标准相比，只能给出“真实的”几何尺寸。

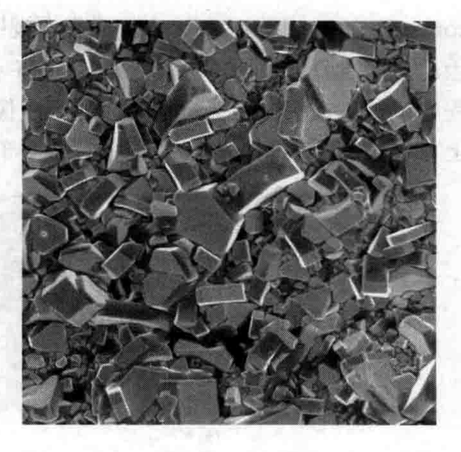


图 15.12 烧结碳化钨表面的 SEM 图像

使用 SEM 检查表面的最大问题是，当图像对人类观察者看起来很自然时，即使用一种熟悉的方式呈现表面几何时，由图像确定实际的尺寸值是不可能的，除非存在约束条件。集成电路测量用于确定横向尺寸，但即使是在这些情形下，边缘的物理轮廓和信号剖面电压间的关系也不确定（这种关系高度依赖于所用电压、被成像材料、检测器类型和位置等）。质量控制中进行一致性而非绝对精度测量很重要，但由这些图像提取 z 方向的测量值很困难（甚至沟槽底部或接触孔洞附近的点都看不见）。

组合表面的两幅（或多幅）不同视图来测量高度的立体成像，原理上与从航拍图像生成地形等高线图相同（Wong，1980；Wang，1990；Matthews，2008）。取决于物体的尺度，使用 SEM、数码相机或其他成像设备可得到这些图像。但很难自动地在两幅图像中找到对应的点（Wrobel，1991；Heipke，1992；Barnard & Thompson，1980；Zhou & Dorrer，1994；Abbasi-Dezfouli & Freeman，1994；Park et al.，2005；Raspanti et al.，2005；Tang et al.，2005；Adhyapak et al.，2007），即使仔细控制成像条件并测量角度，图像的垂向分辨率通常也远差于横向分辨率。这是因为两幅视图间的倾角 δ 通常必须较小（7°～10°），以防隐藏一个视图中的某些点，且角度以 $1/\sin(\delta)$ 进入计算。横向维度的精度会被角度的任何不确定性放大，进而限制最终结果的精度。

当相同的点位于两幅图像中时（要么人为选取，要么通过软件选取），测量各个点间的高度差原理上很简单。如第 1 章所示，应用三角测量法，可由点的视差或偏移得到高度。但在生成整个表面的高度图时，需要匹配大量的点，且实际操作时要求自动化。用于此目的的两种方法是基于区域的方法和基于特征的方法。基于区域的匹配法采用互相关（在空间域和频率域中均可执行）来在第二幅图像上找到和第一幅图像中每个区域最匹配的区域。如果用于匹配的区域的宽度为 5～10 个像素，那么匹配操作可以产生亚像素位置值，但是会限制可匹配点的间距。两幅图像间区域的可见性和对比度变化、存在的镜面反射或产生多个匹配的重复结构，为这种方法带来了许多问题。

特征匹配检测每幅图像中具有某些特性（如最大局部方差或熵，或较高的局部亮度梯度）的位置。这些位置然后再次与另一幅图像中的类似位置匹配。这种方法通常要好一些，但在几百万个像素中，通常只能匹配几千个位置，因此中间位置只能进行内插。对于这两种方法，保持点从左到右的顺序且知道倾向以便匹配仅出现在总图像区域的某个较小范围内，可为实际实现提供巨大的帮助。

第 14 章中的图 14.9 给出一个典型的结果，其中通过互相关已匹配了两幅立体图像，用右图中的可能点来测试左图中的每个点，找到了最佳匹配。匹配处的水平偏移（不一致或视差）即是该点高度的测度。通常会出现误匹配点或未匹配点，使用中值滤波可消除这些点。然后使用计算出的高度值即可进行测量或可视化。为融合通常源于 SEM 的立体像对，人们开发出了许多专用软件（www.Alicona.

com；Ponz et al.，2006）来生成表面视图并进行测量（见图 15.13）。由于扫描电子显微的景深较大，因此常用于获取表面的立体像对。SEM 也用于生成表面成分的 X 射线图，详见下面的讨论。Alicona 程序可校正镜头焦距（短焦距会导致图像的其他变形），并能容纳两幅以上更高精度的图像。在显示器上可自由旋转表面视图，并能执行各种剖面和面积测量。

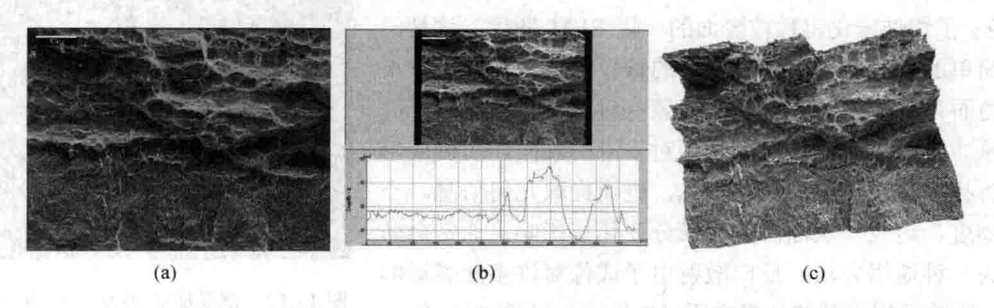


图 15.13 SEM 立体像对所计算表面的可视化与测量：(a)金属断裂表面立体像对的浅浮雕展示；(b)使用 Alicona 软件渲染立体像对所计算断裂表面后的效果；(c)沿断裂表面测量的高度剖面

替代使用电子、光或其他射线来形成一幅表面图像，由生成的散射或衍射图样也可得到有用信息。X 射线图样包含有关于晶体结构或非晶状层的数据。电子散射图样包含有结晶排列的信息和材料中的局部应力（显示为图样中可测量的偏移或线条的加宽）。可见光散射图样包含有表面方向分布的信息（Zhu & Zhang，2005）。这些图样实际上是表面高度轮廓的傅里叶变换，因此按分开表单或图形（零件的大尺度几何信息）、波纹（中等尺度的图形偏离）、纹理或粗糙度（关于表面的小尺度细节）的方式对它们进行测量是研究表面高度的直接方法。在许多情形下，反射光的散射也可与表面的期望用途直接相关。使用处理常规图像的相同方法，可处理和测量衍射图样的图像，但此处不涉及它们的解释。

15.5 表面成分成像

探测表面成分的技术有多种。本书中给出了许多能生成此类图像的仪器。最常的仪器有使用 X 射线检测器的 SEM、使用二次离子质谱（SIMS）的离子显微镜或微探针和傅里叶变换红外（FTIR）光谱学。也可使用许多其他的工具，特别是在测量涂层的厚度和成分时。这些工具包括粒子的反向散射，它在表面下深达几微米处对样本的成分和密度进行采样。相当专业的是声学显微镜，它对涂层和衬底之间的脱胶较为敏感。千兆赫兹声波的波长类似于可见光的波长，因此可用来对（使用其他信号难以检测的）表面和近表面结构成像。表面波会被裂缝强烈反射，体波则会被孔隙的表面强烈反射。测量材料中的声速，可确定弹性模量和其他物理性质。

椭率计量方法利用了这样一个事实，即对于许多类型的薄层介质涂层，偏振光平面会在偏振光通过涂层时旋转。测量该旋转即可得出涂层厚度的高精度值，且在使用不同波长的光（或使用分光仪扫描整个范围的波长）时，也可揭示涂层内容结构的细节。但这种方法主要用于测量相对较大的点，且不会生成表面的图像。类似地，另一种点分析技术直接在表面的一个选取点处使用激光束和光学显微镜来气化表面凹点（通常宽度和深度均为几微米）处的材料，进而用质谱仪对原子和分子碎片称重。

离子显微镜的类型有多种。许多类型的离子显微镜可生成表面的元素组成图，或材料中不同深度的一系列此类图像。入射离子束一次一个点（离子微探针）或一次整个表面地敲开样本最外层的原子。这些原子被电离后，加速到质谱仪中，质谱仪则按照原子的质荷比将它们分开，识别出特定的元素和同位素以及分子片段。检测器或检测器阵列随后将生成一幅图像。这通常表示一种选定元素在整个成像区域的空间，横向分辨率约为 1 微米（在离子微探针情形下，它取决于入射光束的直径；而在离子

显微镜情形下，它取决于离子光学的分辨率），而深度分辨率为 1 个原子层。移除各层时，将分光仪从一种元素切换到另一种元素，就产生了材料结构的完整数据集。

表面成分映射对于淀积涂层的检查和污染的识别尤其重要。最常见的映射方法使用光栅扫描电子束来由原子生成 X 射线，然后检测 X 射线。X 射线独特的能量和波长可识别元素，而根据生成 X 射线的物理原理进行计算，可求出 X 射线的数量。电子的范围将横向和垂向分辨率限制在 1 微米左右。使用了几种不同类型的 X 射线光谱仪；衍射或波长色散型光谱仪每次测量一种元素的 X 射线，但具有较好的追踪元素灵敏度，而更常见的能量色散型光谱仪可同时测量所有的元素，但检测能力较差。这些光谱仪通常会同时生成各种元素的“点图”，如第 1 章中的图 1.31 所示，它们只是近似地勾画出了包含有这些元素的区域，必须处理和合并这些区域。其他信号（如俄歇电子）来自靠近聚焦电子束入口点附近的较小区域，它们具有更好的空间分辨率和深度分辨率，但因为信噪比较差，因此不能很好地检测微量元素。

使用红外光谱可识别涂料和污染的分子，此时激励分子的各种振动模式来产生特性谱峰。这种方法最适合于分析有机材料和涂料，例如塑料。这种方法的空间分辨率因所用光学器件而限制为几微米。

表面成分图通常被显示、处理和分析，此时将信号强度（近似正比于浓度）当成距离图像来处理，并将它形象地解释为一个表面，表面的高度表示浓度，如图 15.14 所示。这些图像通常包含有多个信息通道，每个通道代表不同的元素，它们的不同颜色显示如第 1 章中所示。也可使用第 5 章中介绍的主成分分析来描绘出现的各个相位。

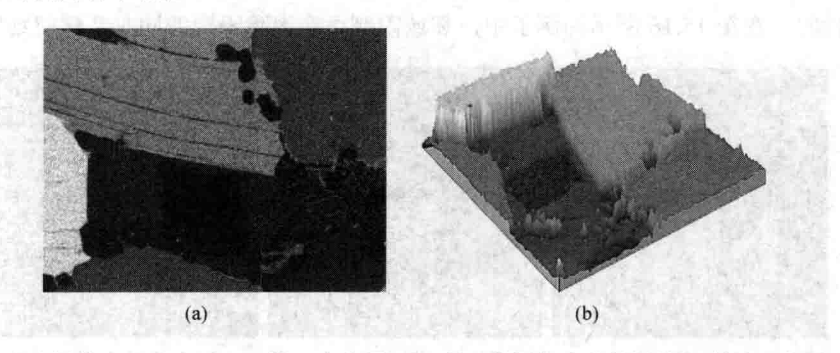


图 15.14　将成分图像表示为表面：(a)第 1 章中图 1.32 所示数据集生成的铝的 X 射线“点图”；(b)表面的等距视图，其中高度代表亮度且近似为元素浓度

15.6　距离图像处理

上述各种方法生成的表面高度数据通常记录为 8 位或 16 位灰度图像，少数情形下也记录为实数阵列。每个像素都有一个值，该值表示了相应表面位置的物理高度或成分。由于多数技术是垂直向下来看表面的，因此数据是单一值，且只表示表面的最高点，因此不能显示下陷和桥梁。如图 15.15 所示，SEM 则是一个例外，它显示了一个复杂的聚合表面，表面上带有下陷和桥梁，而距离图像并不能揭示出这些内容。使用通常存储在文件头中的比例数据，可将距离图像中每个像素的整型数据转换为合适单位（纳米、微米等）的高度值。遗憾的是，数据不存在标准格式，不仅每家制造商都有其自己的标准，而且不同的仪器也会有多种数据格式。

读取这些不同格式的文件并将其存储为某种标准格式（如 TIFF 文件），需要有专门的程序。只要用户能指定（或推断）所需的格式信息，一些图像处理和显示程序就可读取各种格式的数据阵列。这些格式信息至少应包含文件头的长度、特定尺度或其他信息、阵列的维数、数据是按行还是按列存储、

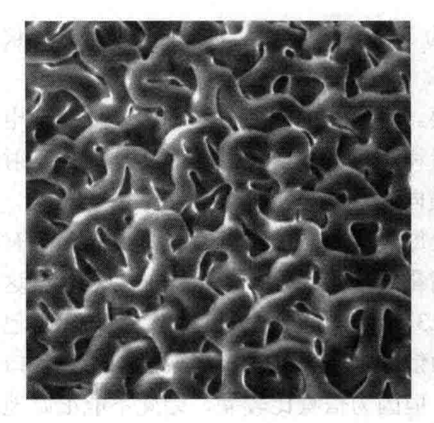


图 15.15　聚合物复杂表面的扫描电子显微镜图像，其中包含了下陷、孔洞和桥梁

数据格式（字节、整型、长整型、实型等，以及字节顺序是低字优先还是高字优先）。

数据可用后，表面图像所需的处理类型很大程度上取决于所用的仪器类型。下面对此举例说明。

干涉显微镜在表面上局部坡度太大的位置，通常会漏失像素，因为在这些位置无法将足够的光返回到光学镜片中来进行测量。事先使用一个非法值或不可能的值来填充阵列中的实测数据，可检测到这些点。具有非法值的任何像素就是一个漏失点，因而必须填充它们，这样做的常见方法是使用中值或平滑滤波。对各个点而言，中值滤波是够用的，但在许多情形下，可能会遗失表面上排列为直线的几个像素（对应于某些阶或脊）。漏失附近像素间的线性或样条内插会产生平滑的结果，但这种结果可能会也可能不会与实际表面相符。

降低或消除距离图像中的噪声时，可采用与其他图像相同的方法：要么使用中值滤波器，要么使用某种类型的平均滤波器（如高斯滤波器）。但最好不要对表面图像应用第 4 章中所示的圆形邻域。相反，邻域被限制为同一表面部分的那些像素，不包括不同高度的点或不同坡度表面上的点。在第 4 章中，这被描述为条件平滑，它是自适应邻域滤波的一个例子，类似于平滑地形数据（即距离图）的过程（克里金内插）。在图 15.16 所示的例子中，邻域限制可在消除噪声的同时，保留细线和角点。

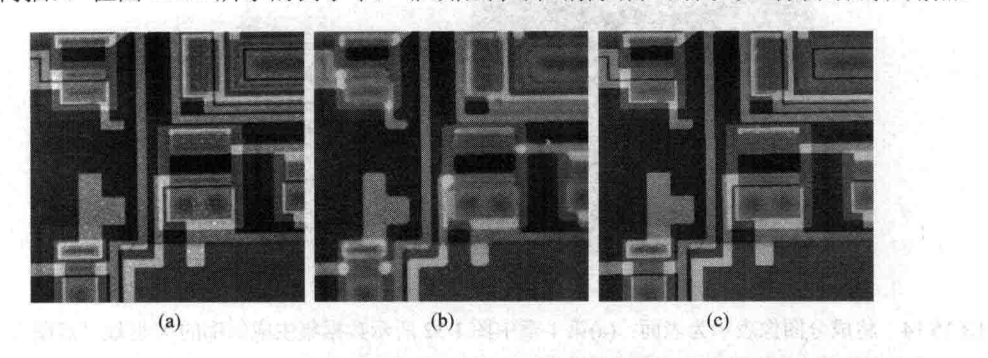


图 15.16　使用一个受限邻域降噪：(a)带噪的原始图像；(b)原始图像经常规中值滤波（7 像素直径）后的结果；(c)原始图像经中值滤波（9 像素直径）和阈值处理后，排除了 40 多个不同于局部表面的灰度值

探针仪器（无论是金刚石刀头半径为几微米的宏观仪器，还是使用巴基管来探测更小横向尺寸的原子力显微镜）还要求各向异性滤波。沿每条线（*X* 方向）的扫描率通常由探针本身的动力学（移动刀头的质量和施加的力）确定，它决定了刀头移过表面并同时与表面接触的最大速度。施加的力太大，则会损毁表面或刀头，施加的力太小，回复力则会使得刀头掠过孔洞或飞过上升斜坡。取决于探针是移动一些传感元件（如干涉仪、电容或电感计，或一束光）还是表面移动到探针的无效位置（经典的 AFM 操作模式）并记录到达压电驱动器的信号，表面高度的输出信号通常是一个电压。必须放大该电压，然后数字化，且在合适的滤波处理过程中使用一个适合于扫描速度的时间常数。因此，这降低了沿每条扫描线的噪声，且不需要再进行后续的降噪数字处理。如下所述，滤波操作可分为与表面形式相关的低频信号和与高频粗糙度值相关的波形，但这只是测量过程的一部分，而不是图像增强。

Y 方向上（从一条线到下一条线）的情形完全不同。在光栅扫描中，两条连续扫描线之间的时间，允许改变机械和电子元件。多数系统仅在一个方向上扫描（以便降低滞后问题），且在探针抬起而不与

样本接触时，进行回扫。当这些方法的垂向分辨率小于 1 纳米时，很难将探针重定位到完全相同的值。因此后续扫描线彼此会垂直或水平偏移。由于眼睛对大范围亮度的突变非常敏感，因此会生成具有明显水平条纹的图像。缓解这一问题的最初方法是调整每条线，以便平均值与前一条线的平均值相同。这种方法并不可取，例如，这意味着如果图像区域中某处存在上升峰或下陷，那么特征周围的背景会偏移，进而产生错误数据，甚至表面的不正确视图，如图 15.17 所示。

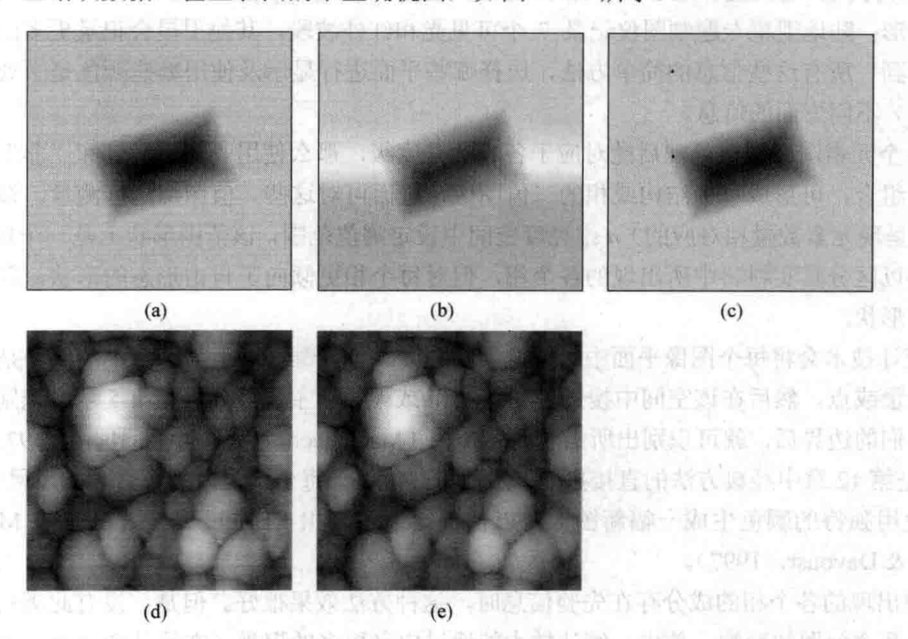


图 15.17　AFM 扫描图像的降噪：(a)硅中蚀刻凹陷原始图像中的线偏移；(b)调整每条线的均值来进行校正后，得到的伪影；(c)使用中值改进后的结果；(d)化学淀积表面的原始图像；(e)校正线偏移后的结果

对于线间调整，有一种比均值或平均值更好的解决方案。理想的解决方案是对齐数据连续线的众数值。假设表面主要由叠加了若干粗糙度的背景电平和关于该平面上下起伏的感兴趣特征组成，不能鲁棒地确定表面高度值。但对于任何分布，中值更接近于众数而非均值。正如与均值相比中值更适合滤除像素阵列中的噪声那样，中值可调整光栅扫描探针图像中的扫描线。图 15.17 显示了应用到典型表面图像的这种校正方法。

通过维护仪器即可避免但在某种程度上可在软件中校正的另一种 AFM 图像伪影，是由图像对针头形状去卷积。如第 6 章中所示，如果可以测量一幅图像的点扩散函数，那么图像的傅里叶变换除以点扩散函数的变换，可消除许多模糊或分辨率损失，因此逆变换会生成一幅改进的图像。实际的 AFM 针头远远不够完美，它会展现出与理想对称点的各种偏离。在已知伪影（如圆盘）上扫描针头（Jarausch et al.，1996；Keller & Franke，1991，1993；Markiewicz & Goh，1994，1995；Villarrubia，1994，1996，1997），可计算出针头的形状并去卷积。实际使用中，针头的寿命很短且针头各不相同，因此需要频繁地进行重校准。但针头去卷积仅在最高放大率和最好分辨率水平下才是有效的。

15.7　成分图处理

由离子质谱仪、X 射线能谱仪或其他方法得到的多数表面成分图，会受低信号电平和随机噪声的影响。第 4 章表明，这种影响原理上可通过收集更多的数据来补偿，但这种方法成本太高，有时甚至

是不可行的，因为分析会更改或消耗表面。因此，通常使用像加权平滑或中值滤波这样的降噪方法。

当得到（显示不同元素或其他化学数据的空间分布的）同一区域的多幅图像时，就需要找到一些方法来显示它们，并将信息呈现给观察者，同时要找到一些方法来描绘和区分所出现的各相。将多幅图像合并为彩色平面就是这样的一种方法，如第 1 章中的图 1.32 所示。但是，由于显示硬件（和人类视觉）的工作方式，这仅能分配三个平面（红、绿和蓝），而图像数量要远多于此。这种情形类似于遥感图像的情形；陆地卫星专题制图仪记录 7 个可见光和红外波段，其他卫星会记录更多的波段。不存在一次“看到”所有这些信息的简单方法，选择哪些平面进行显示及使用哪些颜色是主观的，且能揭示（或隐藏）不同方面的信息。

如果各个元素图经阈值处理后能对应于各相的亮度级，那么使用“与”运算和“非”运算的这些平面的布尔组合，可形成每个结构或相的二值图像，进而可对这些二值图像进行测量。该过程相当于在一个（与呈现元素数量相对应的）n 维亮度空间中设定阈值范围，该范围形状上是一个矩形棱柱。该过程通常足以区分真实材料中所出现的各个相，但对每个相更倾向于自由形态的形状，即对应于亮度自然变化的形状。

有些统计技术会将每个图像平面中每个像素的亮度画为 n 维（对于第 1 章中的例子是 9 维）空间中的一个向量或点，然后在该空间中搜索各个主轴和/或聚类。主成分分析见第 5 章。识别出聚类并定义了围绕它们的边界后，就可识别出所出现的各个相（MacQueen，1968；Anderberg，1973；Hartigan，1975）。这是第 12 章中经典方法的直接拓展。然后将其值位于每个聚类内的像素分类为属于相应的相，对每个类使用独特的颜色生成一幅新图像，以便描绘出各个相（Bright et al.，1988；Mott，1995；DeMandolx & Davoust，1997）。

当期望出现的各个相的成分存在先验信息时，这种方法效果很好。但是，没有此类信息时进行聚类检测会受几个问题的影响。首先，统计技术能通过定义更多的聚类（在统计意义上）更好地分割空间，因此，除非相聚类的数量是已知的，否则结果就值得怀疑。其次，代表结构中很小体积百分比的少数相的聚类，可仅由一小部分点来表示。尽管这些相可能非常重要（例如对于材料的特性和矿石的经济性），但在 n 维空间图上很难定义和检测它们。在跨越各个主要相的边界的背景像素点中，很容易忽视这些相。例如，聚类检测方法更可能基于成分的较小梯度或亮度的统计变化来分割各个相区域，而不是识别出现的少数几个重要相区域。

15.8 数据呈现与可视化

某些类型的表面测量仪器会产生具有大范围分辨率比的数据阵列。换言之，对高度或其他表面特征数据编码的比特数非常大。多数图像处理和显示程序足以应对 256 个灰度级，但即使是这一灰度级，也超出了人类视觉在计算机屏幕上区分它们的能力。1 毫米范围上 1 纳米的分辨率（这通常是高精度探针或干涉仪器的分辨率）会导致 100 万个灰度级（20 位），远远超过了显示或感知的能力。因此，显示程序必须选择整个范围的一部分进行显示，或者压缩数据来显示整个范围。处理可提供帮助，例如，处理局部斜率或导数而非绝对值，但在解释时要求用户应有一定的经验。

来自表面测量工具的大动态范围数据，只能显示在计算机屏幕上，代价是小但可测量的值显示为亮度上不可区分的小差异。在第 5 章中介绍的邻域直方图均衡是可视化这种小变化的一种特别有效的方法，它通常与灰尘和缺陷相关联。图 15.18 显示了本章中介绍的一个例子，即硬币上字母浮雕表面的距离图像。较低硬币表面和凸起字母上的划痕最初都检测不出，但在应用局部均衡过程后，就看得非常清楚，因此它放大了较小的局部差异，同时抑制了较大尺度的差异。当然，对于距离图像，也可使用前几章中讨论的其他处理技术。

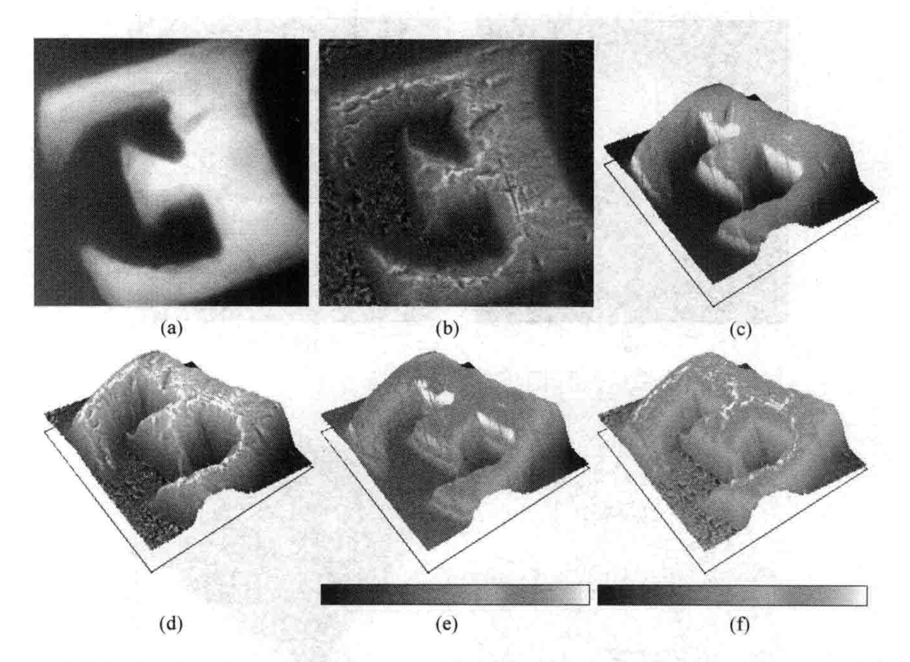


图 15.18　表面标记可见性的增强：(a)硬币表面上字母的原始距离图像；(b)图(a)应用自适应直方图均匀后的结果；(c)图(a)的逼真表面渲染；(d)图(a)中高度信息和图(b)中表面亮度值的渲染；(e)对图(c)应用伪彩色查找表后的结果；(f)对图(d)应用同一颜色表后的结果

距离图像［其中每个像素处的灰度值表示该点的高度（或某些其他表面参数）］包含数据阵列中的所有原始信息。即使显示可容纳这一范围（例如，降低所获取数据的分辨率，或使用局部均衡技术或高通滤波器），结果图像对人类观察者而言也是不熟悉的，因此要求解释人员具有丰富的经验。使用伪彩色增加分辨一些小变化的能力，会使得结果对观察者而言更加陌生，如图 15.18(e)和(f)所示。

等值线图会画出等距线，它们与地球表面的地形图完全相同。许多人对此非常熟悉，因此可读懂这些图，因为沿这些等值线，他们可轻易地识别出半岛和山谷的形状，并识别出具有相同高度的那些点。当然，等值线图中也消除了大量的信息（表面上所有其他像素的高度数据），但这是使得解释更容易的简化的必然结果。图 15.19 显示了硬币灰度距离图像，其中一幅已着色，另一幅已约减为少量的等值线（也已着色，以便更容易区分它们的高度值）。

当细节的横向尺度很小时，或表面呈各向异性时，等值线图基本不能提供可视信息，如图 15.20 所示。此时，各条线紧靠在一起而难以区分，且这些线不能很好地将表面上的不同区域连接到一起。只要等值线因为细节和陡坡出现而靠近时，减少等值线的数量或对这些等值线着色，就可让我们更清楚地观看图形。

表面起伏的视觉效果可通过处理来改善。方向导数可创建一种沿边缘和梯度具有亮和暗对比度的“浮雕”外观。如第 5 章中介绍的那样，形如

$$\begin{bmatrix} -1 & -1 & 0 \\ -1 & \mathbf{0} & +1 \\ 0 & +1 & +1 \end{bmatrix}$$

的权重卷积核从左上角产生照亮表面的效果，并产生人眼解释为起伏的阴影效果。方向通常应从顶部开始，如果亮边缘出现在边缘的底部，人类视觉系统（习惯于光照来自上方）就会反过来解释数据，即将山峰感知为山谷，或将山谷感知为山峰。

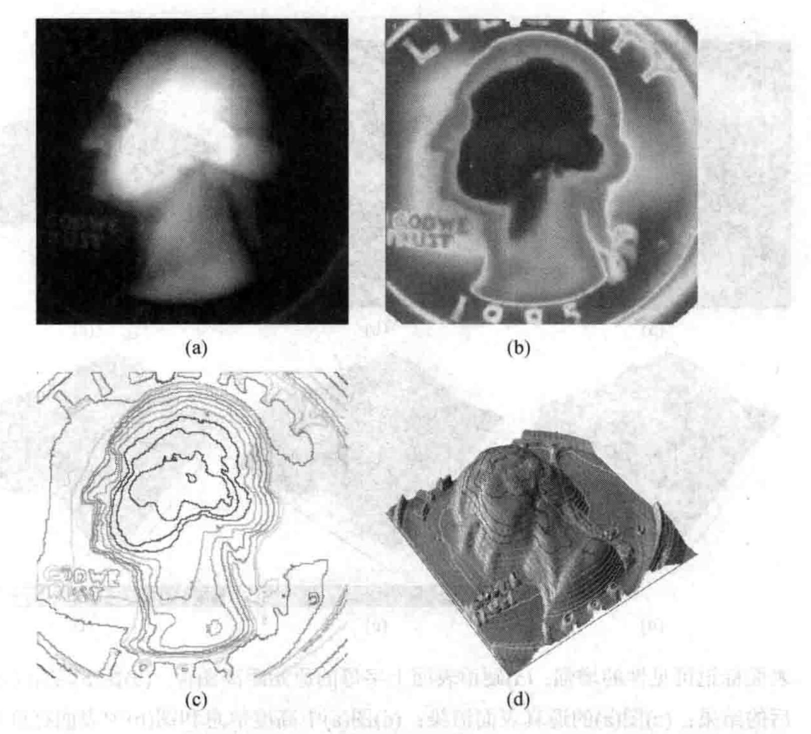


图 15.19 硬币的距离图像：(a)灰度距离图像；(b)着色距离图像；(c)具有 10 条等距线的着色等值线图；(d)渲染后的等距图像，其上叠加了彩色等值线

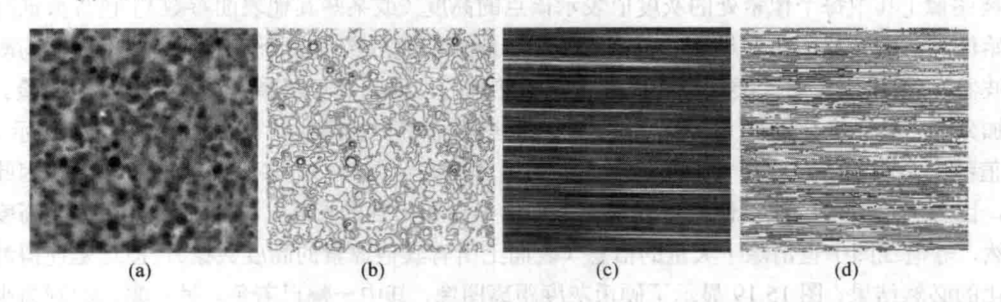


图 15.20 (a)注模聚合物表面的距离图像；(b)注模聚合物表面的等值线图（有 5 条等高线）；(c)研磨金属表面的距离图像；(d)研磨金属表面的等值线图（有 5 条等高线）

导数图像显示了细节，但隐藏了数据中的所有高度变化。组合灰度距离图像和导数图像，可缓解这一问题。将两幅图像相加（在核中将中心值从 0 改为 1）即可实现此目的，但将两幅图像相乘，可得到与视觉感知表面上的纹理更为对应的结果。结果如图 15.21 显示。使用颜色可构建另一种形式的结果。在图 15.21(d)中，使用了一个色调、饱和度和亮度（HSI）模型，将每个像素的高度分配给色调，将导数分配给亮度（饱和度设为 50%）。阴影产生了起伏效果，而颜色则向眼睛告知了整个高度值。

如图 15.22 所示，这些结果可与表面的真实渲染结果相比，真实渲染结果的获取方式如下：使用相邻像素组成的每个三角形作为小面，并计算来自固定位置的光源的反射光。在此类计算中，表面可给出各种反射特性，要么更多散射，要么更多镜面反射。在所示的例子中，用到了一个完整的射线追踪和 Phong 着色。后一种方法会根据相邻三角面间的角度变化来改变跨越每个三角面的阴影，并产生非常平滑和真实的渲染表面，就如 CAD 工作站中所用的渲染表面那样。在亮度通道中对反射率进行灰度渲染，在色调通道中对高度进行灰度渲染，也可对渲染后的图像着色，如图 15.23 所示。

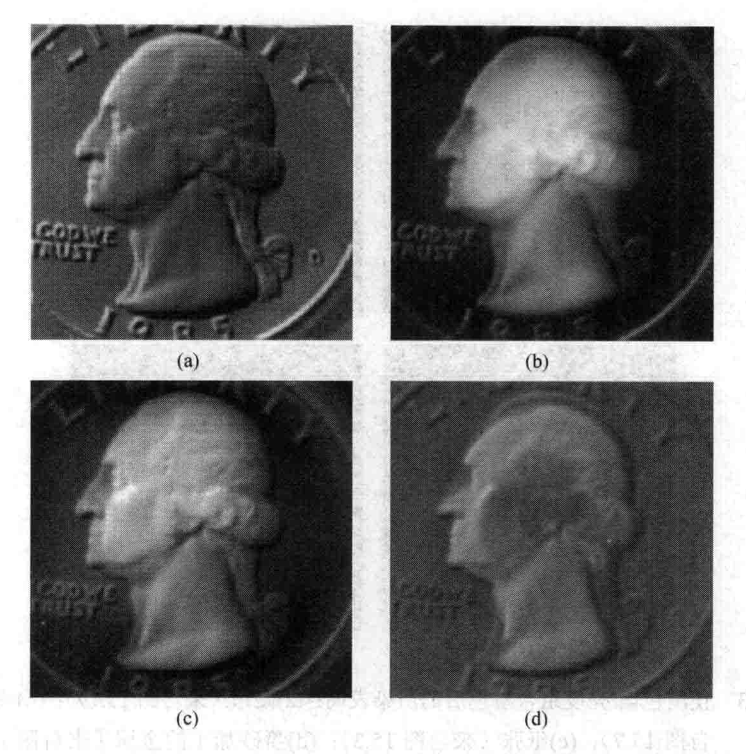


图 15.21　图 15.19 中硬币图像的增强：(a)方向导数；(b)原始灰度值中添加导数后的效果；(c)导数乘以原始高度值的结果；(d)高度在色调通道中、导数在亮度通道中的彩色图像

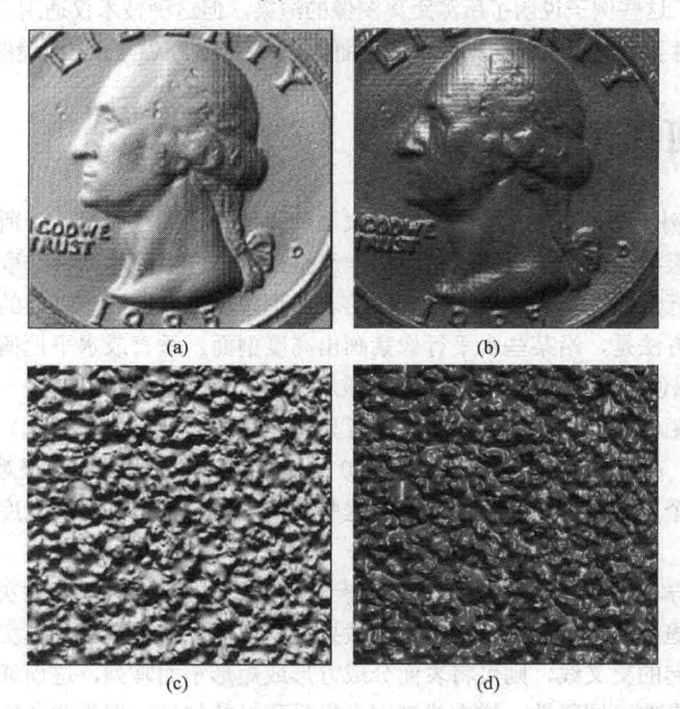


图 15.22　图 15.19 中硬币和图 15.20 中聚合物的渲染表面，它们分别将表面视为漫射散射体（熟石膏）和镜面反射体（发光金属）：(a)硬币，漫射；(b)硬币，镜面反射；(c)聚合物，漫射；(d)聚合物，镜面反射。硬币图像中清晰的网格图案是获取数据所用扫描仪导致的伪影，在这种显示方式下变得更为清晰

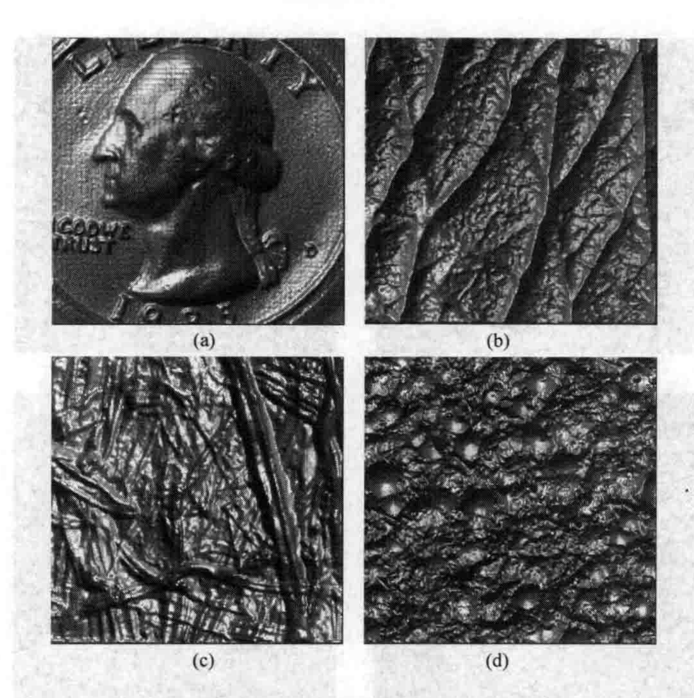


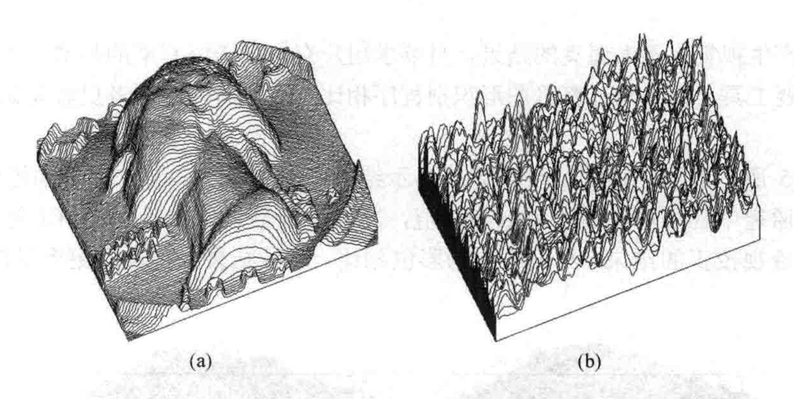
图 15.23 使用色调/亮度组合着色后的渲染表面：(a)硬币（来自图 15.19）；(b)皮肤（来自图 15.7）；(c)纸张（来自图 15.3）；(d)喷砂加工的金属［来自图 15.1(d)］

改变表面的光照和阴影，是检查表面的一种强大的可视化工具，它充分利用了人们根据真实经验来解释表面图像的能力。这些例子说明了所测距离图像的渲染，但这种技术仅适用于由立体测量或其他方式计算的距离数据。对于由阴影成形计算得到的表面数据，图 15.9 说明了表面镜面反射增强的效果。

15.9 渲染和可视化

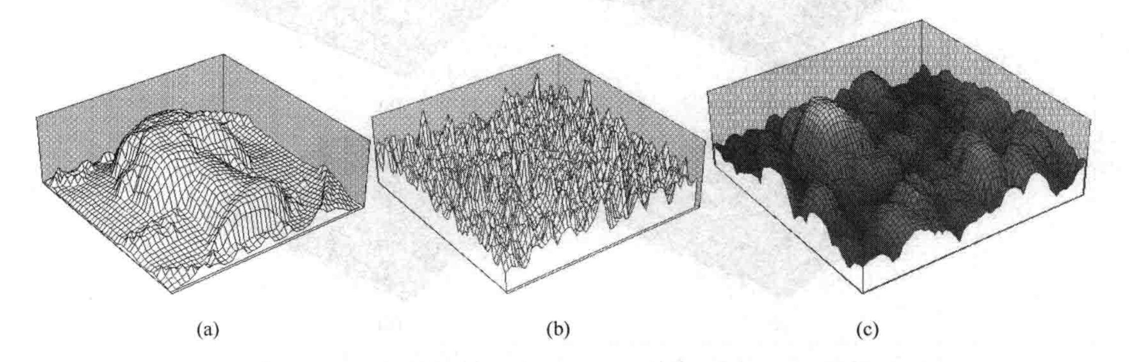
前面给出的多数例子都是直接从上方来观察表面的。这种法向视图显示了所有数据，但人类观察者对这种视图并不是最熟悉的。使用计算机图形学从斜视点（许多程序允许选取斜视点或交互式地旋转斜视点）对表面进行渲染，可生成更容易查看的表面形态，进而检测感兴趣的特征。实现这一任务的最古老和最简单的方法是，沿某些水平行像素画出高度剖面，垂直或水平地偏移每条线来创建表面效果，并跳过一些行以便线间距足够大。图 15.24 显示了两个这样的例子。在第一个例子中，即图 15.19 中的硬币图像，线间距较大，表面坡度很平缓，且只在少数位置隐藏（并删除）了线条。结果是一个非常容易解释的表面。在第二个例子中，即图 15.20 的聚合物图像，这种展示很难解释，因为有许多线条彼此接触，同时整个形态是模糊的。此外，在这些图像中，每条线剖面的宽度是相同的，因此对视图未应用透视效果。

随着计算机图形学的发展，目前的处理能力得到了很大的提升，且显示的灰度级数与颜色越来越多，因此可生成更为逼真的三维视图。添加透视效果也会使得数据看上去更真实，而在连续直线剖面上添加将点连接在一起的交叉线，则可将表面分成方形或矩形小面阵列，进而通过显示第二个方向上的坡度来提升表面形态的可解释性。增大线密度可提供更多的信息，但为避免线条靠得太近，仍需忽略许多线和像素行。按点的高度进行着色，可为深度提供额外的线索。图 15.25 所示的结果，代表了几年前的技术水平，而现在在台式机上就可简单地得到这种结果。诸如 OpenGL 这样的计算机图形学程序包简化了渲染真实表面图像的任务。



(a) (b)

图 15.24　图 15.19 所示硬币图像和图 15.20 所示聚合物图像的等距线剖面图

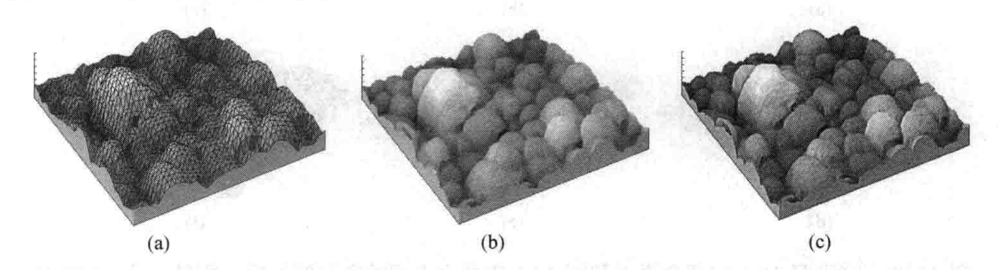


(a) (b) (c)

图 15.25　图 15.19 所示硬币、图 15.20 所示聚合物和图 15.17 所示沉积表面的网格图

视觉上最逼真的呈现会使用实际的表面渲染来控制表面上每个三角面的亮度。与必须弯曲方形小面来匹配 4 个角点（通常不在同一个平面上）相比，三角形小面处理起来更加简单。三个角点定义一个三角形及该三角面关于视线的方向，而由光源位置可计算出将分配给三角面的亮度。给定时间和计算能力后，就可实现类似于摄影的效果，但这通常需要渲染大量的三角面，因此需要找到更快速的方法。根据三角面绝对高度与导数之积来对每个三角面着色时，可快速渲染出视觉上逼真的结果。

图 15.26 说明了这种方法。在图 15.26(a)中，为显示这一过程，画出了各个三角面的边缘。由于这种方法可以为每个像素及其相邻像素创建三角面，因此可显示数据集的整个分辨率。结果如图 15.26(b)所示。在计算机图形学中（如 CAD 程序中使用的计算机图形学），通常会对各个三角面着色，以便它们与其邻域混合在一起（Gourard 着色或 Phong 着色），而不显示它们交汇处的线条。尽管这种方法对 CAD 渲染中遇到的少数较大三角面很重要，但对对应于每个像素的小三角面则没有必要，因为这些三角面在显示结果中仅涵盖 1 或 2 个像素。



(a) (b) (c)

图 15.26　使用三角面渲染的沉积表面，其中三角面的亮度是其高度和坡度的乘积：(a)6 像素宽的较大三角面；(b)连接每个像素及其最近邻像素的三面角；(c)对各个三角面着色后的效果

将这种方法与使用高度来对三角面着色或使用伪彩色来表示高度的简单显示过程相比较，是有用

的。这些方法会产生视觉上不太逼真的结果，且要求用户有一双训练有素的眼睛。人类视觉是检查表面图像的一种重要工具，因为与计算机图形识别程序相比，有经验的观察者更容易识别缺陷或其他感兴趣的特征。

使用图 15.25 所示的同样过程，可在这些显示结果中引入颜色，即在色调通道中将高度应用为颜色，而在亮度通道中应用坡度或 Phong 渲染值，同时在屏幕上的合适位置画出每个三角面，进而生成表面几何经透视校正的结果。与仅应用伪彩色相比，这种结果在视觉上更容易解释，如图 15.27 所示。

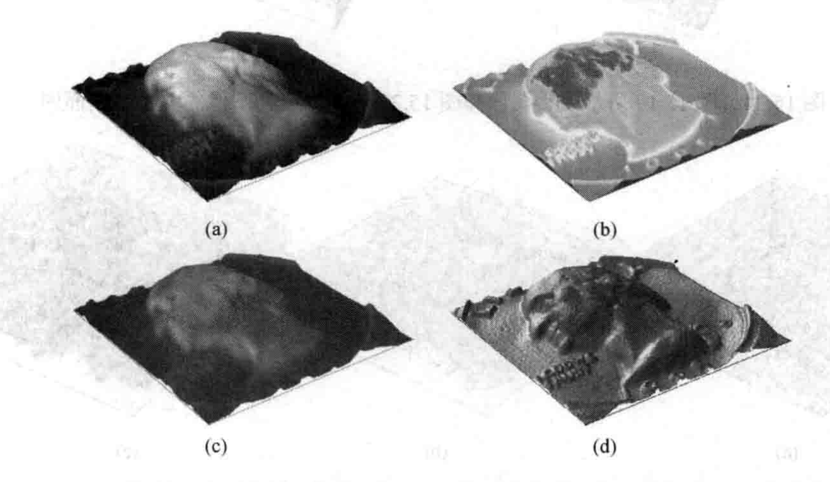


图 15.27　硬币图像的高度数据展示：(a)每个三角面根据其高度和坡度之积着色；(b)对每个三角面分配了一种表示高度的颜色；(c)使用色调表示高度，使用亮度表示坡度；(d)对高度使用色调，对亮度使用 Phong 渲染表面值

对多数用户而言，与简单的灰度距离图像相比，这类显示结果更好地展示了表面的几何形状。图 15.28 给出了图 15.1 和图 15.3 所示相同表面图像的这种显示。即使距离图像包含了所有数据，且经透视校正后的显示结果使其有点模糊，但比较表明后者的外观更逼真，因此在视觉识别特性或缺陷时更有用，但不适用于表面测量。

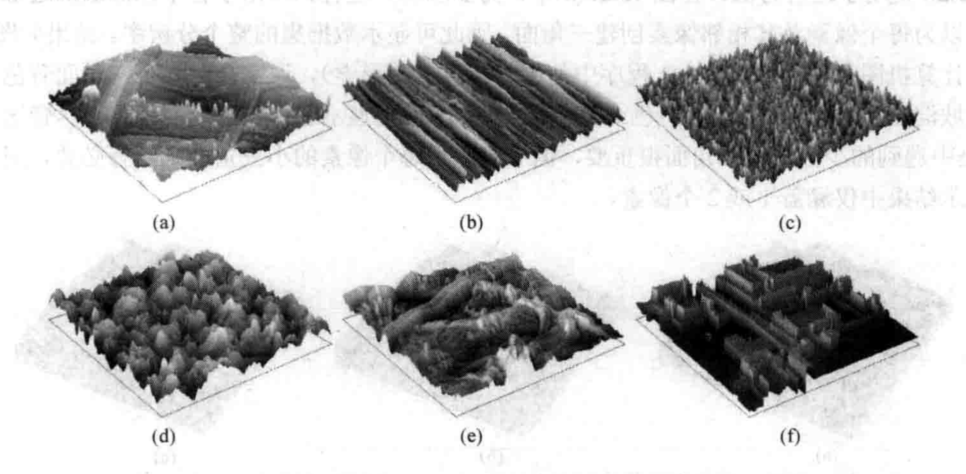


图 15.28　渲染图 15.1 中四个表面和图 15.3 中两个表面的表面数据后，所显示的经透视校正后的结果，其中各个三角面是根据其坡度和高度之积来着色的。这些图形在垂直（z）方向已拉长，以便增强对数据中粗糙度和起伏的感知：(a)加工过的铝；(b)研磨过的钢；(c)蒸气抛光后的铝；(d)喷砂加工的黄铜；(e)纸张；(f)微电子器件

现代计算机图形学也可从不同的视角快速重绘表面视图。其用法有多种。从稍微不同的视角生成一个表面的两幅逼真渲染视图，就可通过人类立体视觉来解释表面的深度。使用从不同视角生成的一系列此类图像，可创建飞过表面的“电影”，现代计算机可实时完成此任务，即操作者操控游戏杆来交互地控制飞行路径。这种效果与立体显示相结合（见图 15.29），就可生成一种虚拟现实世界，在这个世界中可详细查看表面。这幅图像使用双色立体图像演示了这种效果，观看时需要红色/绿色或红色/蓝色眼镜，但也可使用其他显示方式。

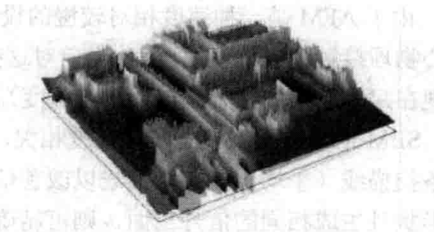


图 15.29　从两个稍微不同视角渲染图 15.28(f) 后的效果，这两个不同视角的表面渲染可合并为一个立体像对（使用红色/绿色或红色/蓝色眼镜观看）

15.10　表面数据分析

人工视觉检查表面几何和成分信息的渲染结果，是检测和识别表面缺陷及其他特定特性的强大工具。但对于许多目的而言，需要对表面进行数值测量，进而控制这些目的，并将表面几何或成分变化与表面的创建和处理历史关联起来，或与其表现行为关系起来。对于这些目的，需要依赖于计算机来处理数据的解析方法，而需要测量什么来为任何给定的需求提供有效的参数并不总是很明朗。

传统的表面测量包括涂层厚度和表面几何的几何特征测量。多数这样的测量值都高度依赖于其他因素，如材料成分（和成分的空间变化）、所用的特殊测量过程、待测区域的大小。多数厚度测量过程和一些高度测量仪器会在几微米的横向距离上进行平均。位于一个区域内最高点的探针仪器由探针针头的大小确定。与多数技术实现的垂向分辨率相比，这要大得多，且会隐藏表面或涂层的重要细节。因此，必须采用抽样策略来确定空间均匀度。

表面的尺寸测量是使用了坐标测量机、探针设备和光学干涉技术的一种应用。在许多情形下，在零件设计时指定了非常精确的尺寸，因此测量不需要对一个区域或一幅图像进行扫描，而只需要将测量工具与零件正确地对齐。但随着尺寸变小，就像微电子器件那样，则需要获取一幅图像来定位待测点。图 15.30 显示了一个例子，它是取自 AFM 扫描线的高度剖面，由该高度剖面，可测量各个阶的宽度和高度。AFM 数据集通常显示为表面高度视图，但定量测量正变得越来越常见（Shuman，2005）。

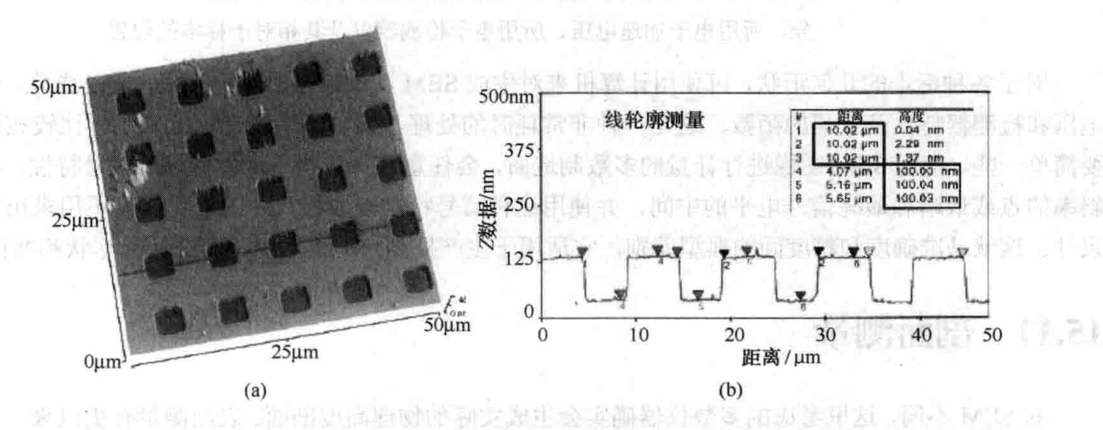


图 15.30　(a)为单线扫描选取位置（红色）的平板测试图案的 AFM 图像；(b)尺寸测量所用的单线扫描。此时，AFM 要求进行定量位置感知（如干涉仪），而非依靠测量送入压电定位器的电压

由于 AFM 是一种速度相对较慢的设备（它很难处理较大的零件或扫描较大的区域），且总是需要担心物理接触会损坏表面，因此当前对这些器件的测量主要由 SEM 完成（较早的设备使用光学显微镜，但现在尺寸已小到光波无法分辨的程度）。图 15.31 显示了硅晶圆上两条光刻胶的 SEM 图像。遗憾的是，SEM 信号并不直接与表面高度相关，而主要取决于局部坡度、粗糙度和成分。因此，很难解释沿一条扫描线（平均了多条扫描线以改善信噪比）的信号剖面，进而求出线宽。若假设这些线具有相同的形状且生成相同的信号剖面，则可精确地求出线距。因此，选择信号的任何可再生特性（峰值或最大坡度等）就可测量线距。而要精确地测量线宽，则必须确定边缘所在的绝对位置。

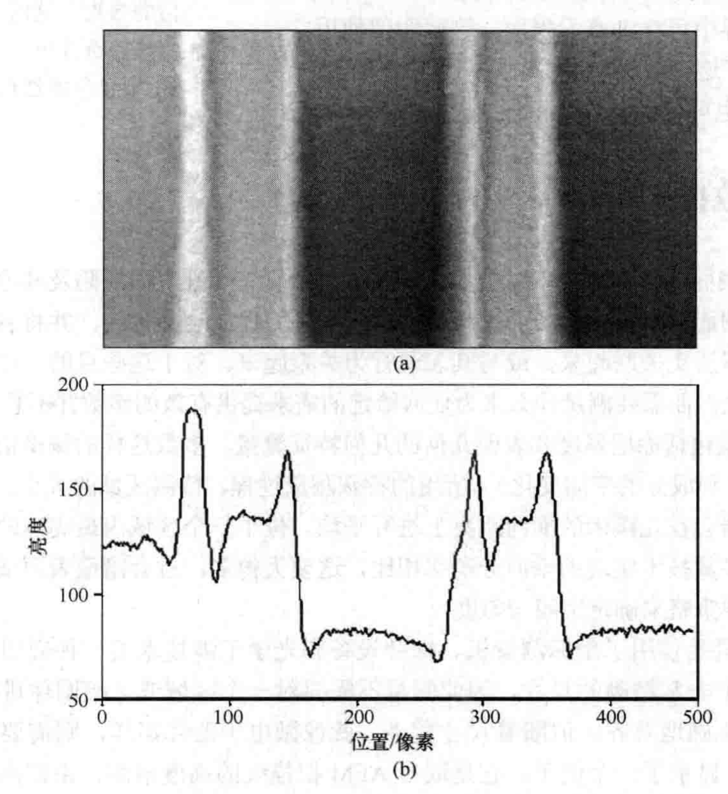


图 15.31　晶圆上两条平行光阻材料的 SEM 图像和横跨它们的信号剖面。物理剖面和信号间的关系取决于这些线条两侧的坡度和粗糙度、材料成分、所用电子加速电压、所用电子检测器以及其相对于标本的位置

对于各种标本的几何形状，可使用计算机来对生成 SEM 图像信号的过程建模，它是成分、电子束电压和检测器特性及布局的函数。这是一种非常耗时的处理，但与制造物理标准来进行比较相比，仍要简单一些。使用 SEM 图像进行计量的多数制造商，会任意选择一些重现性较好的信号特性，如最大斜率的点或最暗和最亮信号电平的中间，并使用这种信号特性来监控尺寸的变化，而不用求出实际的尺寸。这就是准确度和精度间的典型差别，它适用于生产控制，但不适用于新的几何形状和器件。

15.11　剖面测量

和 SEM 不同，这里考虑的多数仪器确实会生成实际的物理高度剖面。表面测量有史以来一直是由这些高度剖面评估的，而不是使用全部二维图像来评估的（因为这种仪器更简单且更便宜，所需要的时间更少，且人们更熟悉这些方法）。对数据滤波后，就可分离剖面中的不同频率范围，它们通常称为

形态、波纹和纹理或粗糙度（见图 15.32）。

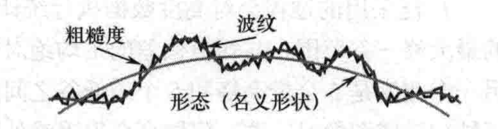


图 15.32　与预期形状的偏差通常分为粗糙度（小尺度）、波纹（中尺度）和形态（大尺度）

形态是整体的粗略几何形状，它一般在工程图中指定，并受设定尺寸控制，由传统的欧几里得几何描述。中频通常称为波纹，而高频则称为纹理或粗糙度。在加工过程中，波纹假定来自机器中的振动或偏转，而粗糙度来自工具和其在材料中遇到的微结构间的局部相互作用。这样的划分具有很大的随意性，它会因零件尺寸的不同而不同。记住，探针本身的形状是一个消除高频的有效滤波器。定义滤波器的截止频率通常设置在 0.25 毫米到几毫米的波长之间，以便将波纹与粗糙度分开。国际标准将它们规定为许多机械工程应用的测量过程的一部分。

采用滤波方式在波纹和形态数据中分开粗糙度，最初是使用电子学中的模拟 RC 滤波器来实现的。在现代数字处理系统中，使用了最小二乘线或弧拟合来删除形态，并使用空间高斯滤波来分开波纹和粗糙度。这种方法可直接推广到区域扫描中，此时也可使用一个等效的高斯滤波器来进行滤波，可对图像的傅里叶变换进行滤波，进而选取期望的频率范围。最常使用平面或某种欧几里得形状（如对应于已知期望形态的柱体或球体）或广义多边形的最小二乘拟合来分开形态数据。图 15.33 显示了形态删除的一个简单例子，其中滚珠轴承（球体）表面的粗糙度看起来很明显，因此在减去背景曲率后很容易测量。然后根据球体的名义形态，即可测量出偏差。

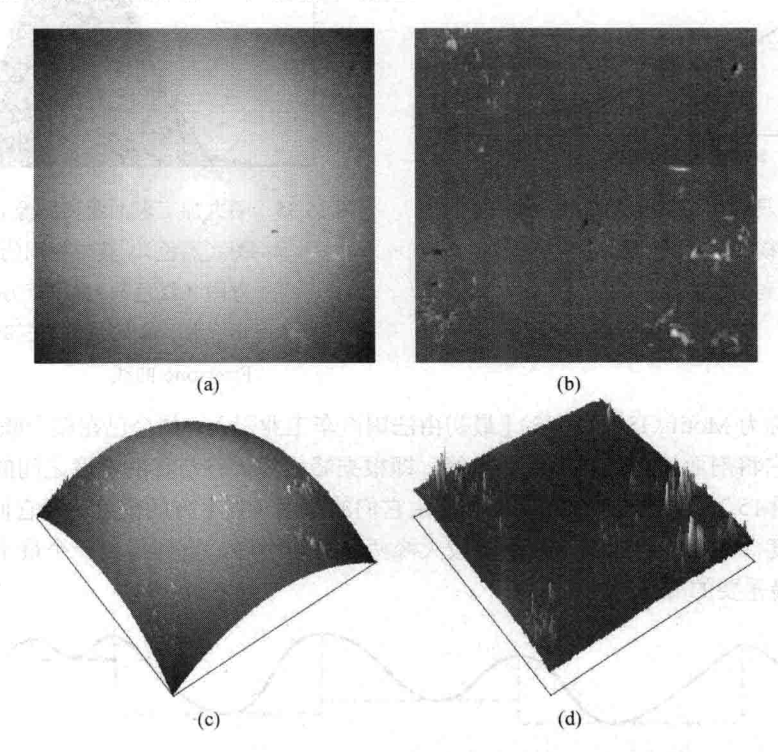


图 15.33　滚珠轴承表面的形态去除：(a)灰度距离图像；(b)减去球形后的距离图像；(c)图(a)的表面渲染；(d)图(b)的表面渲染。去除形态后，数据扩展起来更方便

去除低频分量后，就可由粗糙度剖面求出表面的粗糙度。此时会用到许多测量参数，有些参数已编录到了各种 ISO 或其他国际标准中，有些参数则与特定行业或装备制造商相关（Rosen & Carfoord，1997）。关于仪器和方法的完整回顾，请参阅 Whitehouse（1994），关于分形过程的最新进展，请参阅 Thomas（1999）和 Mainsah et al.（2001）。

广泛采用的过程会对高度数据执行统计分析，而不关心其空间排列方式。这样的例子包括沿剖面的最大峰－谷范围、与均值偏差的平均绝对值（Ra），或高度数据的统计标准差（Rq）。粗糙度大小的另一个测度是 5 个最高峰和 5 个最低谷之间的高度差（Rz），它要求定义峰或谷。如下所述，当应用于区域扫描或图像时，这一问题会变得更难处理。

高度数据的空间分布信息包括一些参数，如剖面中峰的数量和相关长度。相关长度定义为两个连续峰之间的平均距离，或位于某些特定高度（如删除形态和波纹后留下的平均高度线）处的点间的平均距离。相关长度更为通用的定义如图 15.34 所示，它只是由该剖面的傅里叶变换求出的自相关函数的大小。第 6 章中描述的自相关函数（ACF）也很重要，因为对于由大量独立事件（喷丸、研磨、喷粒沉积）产生的表面来说，它的表面形状与产生该表面的“平均事件”的 ACF 相同。

还会用到一些与特殊表面相关的功能参数。Abbott-Firestone 曲线是高度数据的累积直方图（见图 15.35），它会给出删除部分表面、磨损或其他制造步骤（如汽车汽缸衬垫的平顶珩磨）导致的接触面。

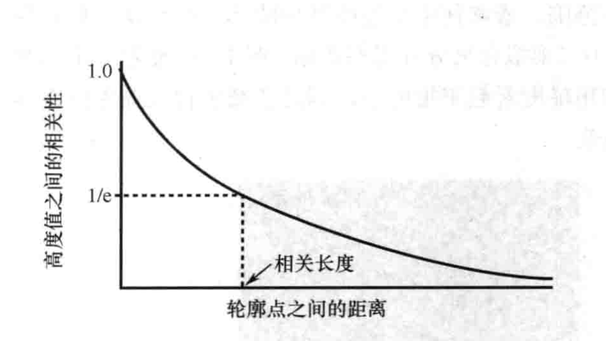


图 15.34　相关图将具有相同高度值的点的概率显示为这些点的横向分离的函数。相关长度定义为图中下降到 1/e 或 36.79%的点

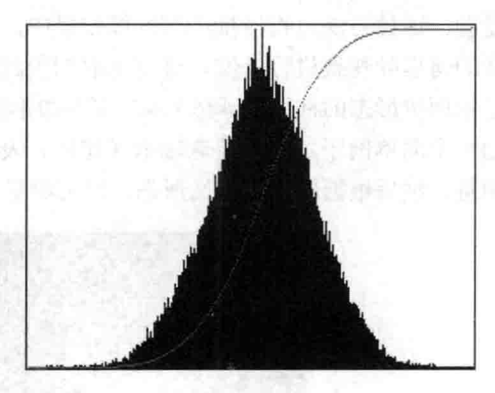


图 15.35　喷丸加工表面距离图像［图 15.1(d)］的直方图（黑色），其中相同的数据显示为累积直方图（红色）。累积直方图表示表面上高度小于水平轴上的值的一部分点，它称为 Abbott-Firestone 曲线

另一种方法称为 Motif（ISO 12085）［最初由法国汽车工业引入，如今已在整个欧洲使用（Dietzsch et al.，1997）］，它将剖面简化到只出现各个峰，即根据峰相对于谷与谷的宽度之间的相对高度来与另一个表面相接。图 15.36 给出这一原理。各个峰由它们的深度（谷上方的高度）及它们的间距表征。然后根据间距和深度合并峰和谷图案，来消除较大峰两侧的较小峰，直到得到一个最小表示，这个最小表示中就只包含最重要的峰。

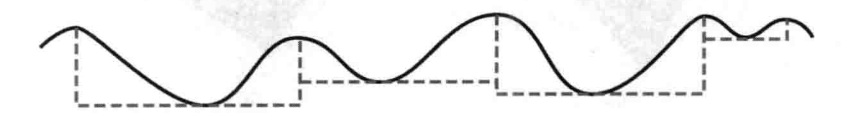


图 15.36　图形合并原理：每个峰和谷图形由深度（谷上方最小峰的高度）和宽度测量。剖面中，与两侧的峰相比要小、宽度小于任意一个截止宽度、深度小于最大深度 60%的图形，将会与其相邻图形合并，进而降低所出现图形的数量

这些方法都会简化原始数据，以便提取一些关键值，但也都存在一些局限性，即它们高度依赖于扫描轮廓的长度和数据点的横向分辨率。这些方法在涉及峰与谷的构成时，定义非常随意，忽略了剖面不经过多数表面峰和谷的最高点或最低点这一事实，且不能与感兴趣的真实物体很好地关联，而关

联则是生成表面及表面在所用服务中的行为的处理过程。这些方法主要适用于特定质量控制应用，在这些应用中，隐藏了参数的真实意义，但可用测量结果的一致性控制操作过程。此外，线剖面的解释很难推广到由高度区域扫描生成的表面图像。

所有剖面方法都受这样一个事实的影响，即多数真实的表面不是各向同性的，而是具有方向性的，方向性源自生成表面的方式、材料本身的特性等。这种所谓的表面“放置”可以是简单的[例如图 15.1(b)所示地面是高度方向性的]，也可以是非常复杂和微妙的。建议测量垂直于主要放置方向的一个剖面，但对于复杂的表面而言，这会丧失表面的许多实际特性。

因为剖面测量产生（或累积）了大量的参数，因此人们正试图使得区域扫描的测量合理化。在得到 ISO 委员会的支持后，伯明翰大学的学者牵头（Stout et al.，1993）提出了一套统计的、空间的和功能性的参数，这些参数将会是未来国际标准的基础。这些参数仍有一些与剖面测量相同的限制，如需要定义峰由什么组成，且强烈依赖于扫描区域的大小和各个点的横向分辨率。此外，它们并未包含一些可能很重要的方法，如地形分析、包络或模式分析和分形几何。但由于它们是描述表面的起点，因此对它们进行考虑是适当的。

15.12 伯明翰测量套件

测量参数共有 4 类：一类处理高度值而不用考虑它们的位置（称为振幅参数），一类处理表面的横向距离（称为空间参数），一类将前两类参数合并在一起（称为混合参数），一类与表面历史和属性直接相关（称为功能性参数）。在每类参数中，只选取少数参数，这些参数与广泛接受的剖面测量参数最为相关。为这些参数建议的符号使用了剖面方法中相同的术语，但 S（表面）取代了剖面标准中所用的 R。

振幅参数是剖面图中所用区域扫描统计测量的简单扩展。例如，Sa 类似于算术平均偏差 Ra。对于区域扫描，它是来自平均平面的高度绝对值的算术平均。保留 Sa 的原因是 Ra 应用的广泛性，反之亦然，因为在计算机出现前设计仪器来测量它是非常简单的。高度点的均方根偏差是一个更为鲁棒的测度，它由高度值分布的标准差 Sq 给出。方差（标准差的平方）是分布的二阶矩。三阶矩和四阶矩分别是偏度和峰度，且它们也用做振幅测量参数，分别称为 Ssk 和 Sku。对于不是双峰的简单分布而言，这三个参数可简洁地统计描述表面高度。

表面高度数据的直方图（如图 15.37 所示的例子）显示了表面高度的整个范围。分布中偏态将此类情形分辨为窄且深的凹槽，而这些凹槽对于分辨汽车发动机中汽缸套上的润滑剂很重要[见图 15.37(c)]。此时，多数表面的高度范围都很窄，而仅覆盖了一小部分区域的凹槽会向下到达低得多的深度。相反方向的偏态对应于这样一个表面，该表面上仅有少数较高的峰或脊。但直方图本身并不包含凹点和谷或峰和脊的空间排列信息。由另一种表面（连续脊上的所有高点聚在一起，或以几千个较小的独立峰的形式分布）也可得到类似的直方图。这两种极端表面的性质将完全不同。

如剖面一样，这些统计振幅测量对采样区域非常敏感。对于多数表面而言，标准差会随观测的点数增加；事实上，对于分形粗糙表面而言，在对数轴上将方差画为测量区域大小的函数，是测量分形维数的一种方法。

对剖面而言，参数 Rz 是 5 个最高峰的平均值和 5 个最低谷的平均值的高度差。对于表面的区域扫描，它可以推广为 Sz，即 10 个最高峰的平均值和 10 个最低谷的平均值的高度差。但它并不是一个纯粹的振幅参数，因为它完全取决于峰和谷的定义。它们不能简单地是表面上的最高点和最低点（也不能是表面图像中的像素），因为这些点可能（且通常）是彼此相邻的，进而表示的只是单个峰和谷。对

于剖面，分开两个高点的一个低点可用来指示单独的峰。这是一种有缺陷的定义，因为明显需要考虑无穷小的不规则性，因此需要为峰间谷的深度和宽度制定某些标准。对于表面的区域扫描，甚至需要更多的此类标准，因为峰（或谷）涵盖了一个区域，且两个或多个局部峰会沿复杂路径（桥）连接，因此会认为是同一个峰的一部分（对于谷而言，反之亦然）。在跟踪这种连接性时，需要考虑像素是与其所有的 8 个相邻像素相接，还是仅与其中的 4 个共边像素相接。

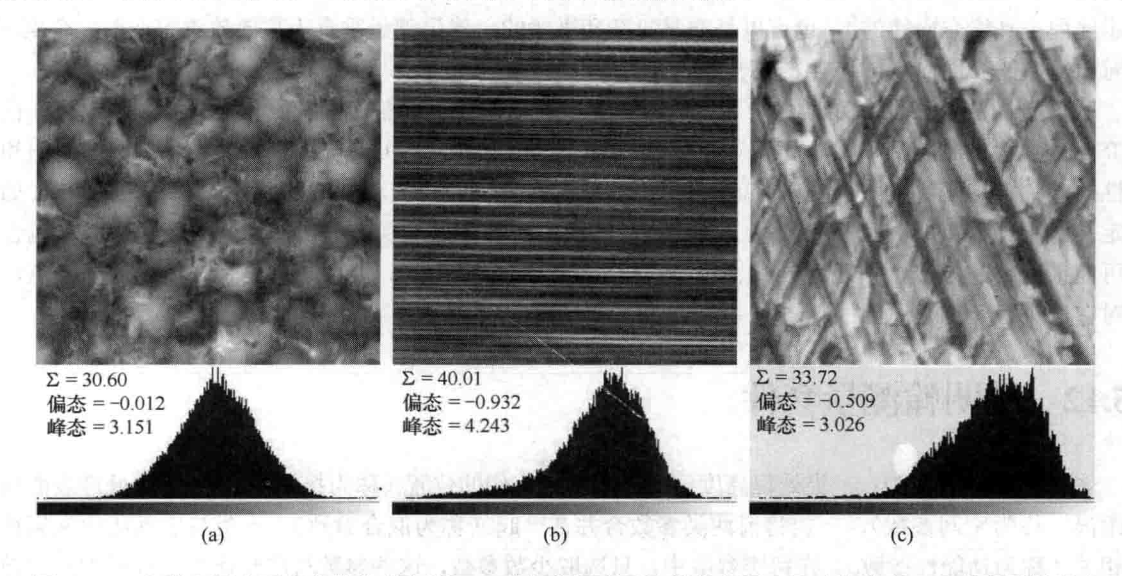


图 15.37　几种金属表面高度值的直方图：(a)喷丸黄铜表面及其直方图，呈对称分布；(b)研磨不锈钢表面及其直方图，由于存在几个深且分开的平行沟槽，因此稍微有点负斜；(c)高原珩磨缸套表面及其直方图，分配润滑剂的较深且相交的沟槽使得直方图更为负斜

在辨识峰和谷时，与只使用参数 Sz 相比，会具有更多的信息。不同表面会有不同形状的峰和谷，峰和谷的大小、形状、方向和间距都包含有重要的特性信息。图 15.38 中给出了几个例子，其中的峰定义为 8 连通的（像素与 8 个相邻像素相接），且要求明显下降到峰高的 80%。谷以同样的方式定义。在这个例子中，找到了 20 个最高峰和最低谷。该方法类似于图像处理中所用的“洪水填充”算法，即从最高的局部最大值开始（并向下进行），并包含向下延伸到 80%极限的所有相接像素，同时检查该峰是否合并到了一个原来已标记的峰中。对于喷丸加工的表面，各个谷轮廓上是相对平滑的，但各个峰则非常不规则。对于研磨表面而言，峰（脊）通常要比谷（缝隙）宽，而对于磨光表面而言，峰非常大而谷非常小。所有这些差异都与理解此类表面的生成方式是一致的，且它们可为其他表面及这些表面的性能提供重要线索（Sacerdotti et al.，2002）。

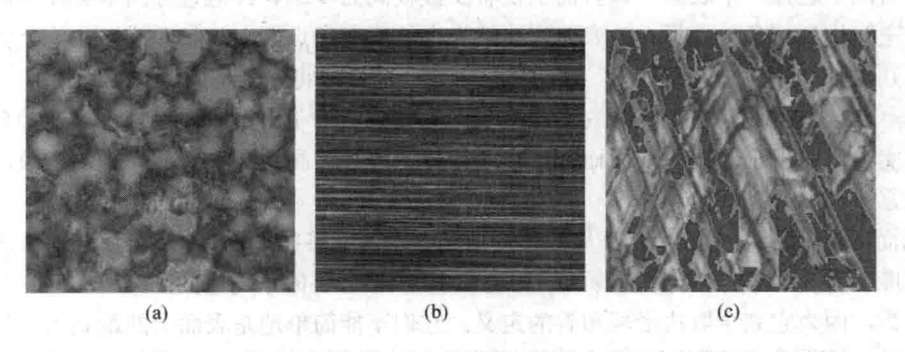


图 15.38　图 15.37 所示表面上的 20 个最高峰（红色）和 20 个最低谷（绿色）

涉及表面上各个峰的另一个参数是单位面积上的峰数，称为 Sds。它同样取决于峰的定义。在许多应用中，关于峰（和谷）的其他信息也很重要。例如，在表面涉及电气或热接触、摩擦和磨损的情形下，或在判断视觉和美学外观的情形下，各个峰间间距的均匀性会很重要。如第 10 章中所示，使用平均最近邻距离来求出是均匀间隔还是聚集的倾向，方法是将该值与平均距离相比较，平均距离是在单位面积上相同数量的点的泊松随机分布。图 15.39 显示了喷射成型的聚合物表面，其中的各个峰稍微有点自避，它们是由冲模的表面抛光、喷射过程中所用的温度和压力、分子重量和聚合物黏度之间的复杂交互形成的。对于随机分布，测量的最近邻距离与预测距离之比为 1.18。这种均匀性与接近于人类视觉空间分辨率极限的两个峰之间的距离相结合，可产生非常令人愉悦的外观。

图 15.39　图 15.20(a)所示模塑聚合物表面上 1 平方毫米区域内的 20 个最高峰和 20 个最低谷。峰根据其高度从紫色到红色着色，而谷从绿色到青色着色。注意包围一个峰的一个谷，且这些峰在形状上要比谷更规则

注意，在该图中，谷的形状与分布与峰的完全不同。在某些其他应用中，人们会对凹点而非峰的密度和均匀度信息感兴趣。一个例子是印刷业中用于保留油墨的图版表面。

具有各向异性的表面或层可由自相关函数导出的空间参数来表征。自相关函数（ACF）的获得方式如下：对傅里叶变换中的复变量的幅度进行平方，同时将相位设为零（这消除了所有的空间位置信息）。反变换会生成 ACF 的二维空间图像。由该函数定义的参数是纹理长宽比 Str、纹理方向 Std 和自相关长度 Sal。

研究图 15.40 中的 ACF 图像有助于理解这些参数。在伯明翰报告中，自相关长度定义为 ACF 幅度下降到 20%的最短距离。对于给出的例子，它是以 20%的亮度级画出的等值线的最小半径。纹理长宽比是最小半径与最大半径之比，而纹理方向是最大半径的方向。对于图 15.40 中的例子，聚合物表面的 ACF［见图 15.20(a)］是各向同性的（表明表面也是各向同性的），因此长宽比是 1，即没有明显的方向。相反，研磨表面［见图 15.1(b)］有一个非常明显的首选方向，该方向在 ACF 中很明显，且可以测量。对于快速切削表面［见图 15.1(a)］而言，纹理更加复杂，因为原始图像显示了两个主要的加工方向。就像定义的那样，由 ACF 可测出 Str、Std 和 Sal 的值，但它们是否包含表征表面的所有信息并不明朗。

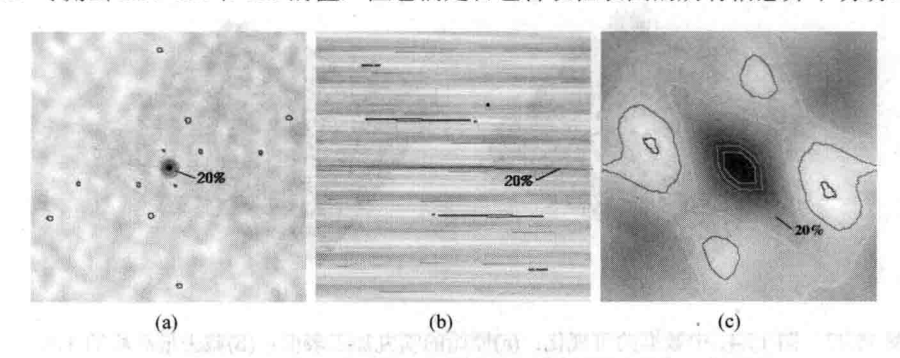


图 15.40　(a)对图 15.20(a)所示聚合物表面计算的自相关函数；(b)对图 15.1(b)所示研磨表面计算的自相关函数；(c)对图 15.1(a)所示的快速切削表面计算的自相关函数。叠加的等值线指出了函数的形状及向下 20%的距离

混合性质涉及高度和横向数据（如前几个参数一样）。SΔq 是表面的均方根斜率，它可由用于生成可视化结果的相同三角形小片过程计算。它的正式定义是，图像在垂直方向和水平方向上的偏差的平方和的均值的平方根，它可由相邻像素间的局部高度差求出。类似地，平均峰曲率 Ssc 与高度的二阶导数相关。当然，它主要取决于峰的定义要有意义并能被人接受。第三个混合性质是实际表面面积与投影面积之比 Sdr，它可通过对组成可视化图形的各个三角面的面积求和得到。

计算机出这些混合性质并不难，但它们完全取决于采样间隔或像素的间隔。改变横向分辨率也会戏剧性地改变参数值，因此它们与测量技术相比，更少与表面相关，故仅适用最受限的比较。基于剖面的许多测量也存在这一问题，但改为基于区域进行测量的目标之一是克服较老方法的某些局限性。事实上，许多工程表面已被证明是分形几何，其实际的表面面积无法确定（随着横向测量分辨率的增大，面积也会增大）。

这是一种很微妙的情形，但此类测量也取决于每个像素处的高度数据是表面的点样本还是整个像素区域的平均。数学应用于抽样情形，此时每个像素处的高度以精确的数学点测量，而不考虑该像素和下一像素间发生的任何事情。事实上，许多测量方法（诸如探针仪器和光学干涉仪）会对整个像素区域执行某些平均测量，并报告该区域的最大值或高度值的加权平均。适合于这些情形的计算不仅会影响混合参数，而且会影响此处描述的所有参数。

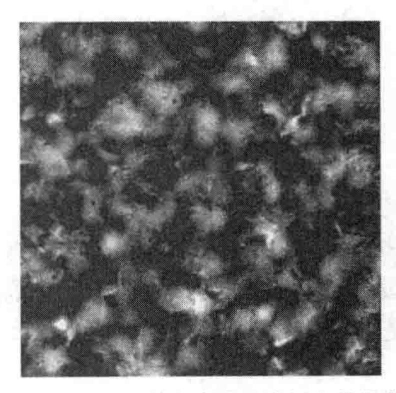


图 15.41 在任何特定高度（本例中是最大值的 31%）对距离图像［本例中是图 15.1(d)中的喷丸加工表面］进行阈值处理后，显示了与一个有着相应磨损量的平面相接的表面区域（缺少弹性或塑性变形）

功能性参数意在将表面几何数据与表面性能的某些方面关联起来，有些参数通常与机械工程应用相关，因为到目前为止，大量使用表面计量的是机械工程领域。一个典型的例子是表面轴承面积比 Stp，在通过磨损删除给定高度的所有峰后，它是与平行于基底的一个平面相接触的部分图像区域（见图 15.41）。该值可直接由图像高度数据的直方图读出。

类似地，处理中删除的体积大小（材料体积比 Smr），可通过对直方图积分或使用累积直方图计算出。空心体积比 Svr 是标本表面内空心空间的体积，用于保留或分配润滑剂。它可由 Abbott-Firestone 曲线测出。图 15.42 显示了与图 15.41 相同的数据，但它显示了相接表面的性质，以及因磨损而删除某些峰后的空心体积（但要假定剩下的表面无变形，且凹点处未填入碎片）。

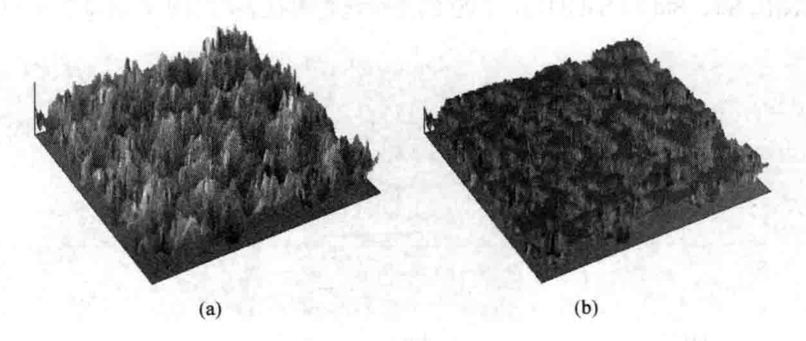


图 15.42 图 15.41 中数据的可视化：(a)原始的喷丸加工表面；(b)截去最高峰的 31%后的同一表面，显示了接触面和空心体积

与这些参数相比，需要提供这些接触面的更多信息。各个触点的大小对于加热和变形而言非常重

要，且空心体积要么完全连通，要么由孤立的小块区域组成，要么是前两者的混合，它们对润滑而言会产生截然不同的结果。人们还提出了其他的功能性参数来处理这些性质和其他方面的表面性能，但它们正成为每个应用领域的专用参数，且要求正确地定义和使用。许多这样的参数都存在缺陷，因为距离图像中的表面高度数据是单值的。在每个像素处记录的高度用于该位置最上面的表面，就如被探针或光学反射等检测的那样。如果磨损删除了某些表面覆盖，那些距离图像中未显示的表面内的底切、洞穴或孔洞可能会变得很重要。

使用前几章中开发出的工具对图像进行处理和分析，可得到距离图像表面特性的许多有趣参数。例如，可以使用最小和最大排序算子（灰度腐蚀和膨胀）来修改图像，进而形成表面包络，创建上面介绍剖面时提及的图案逻辑的二维形式。与缺陷（裂纹、凹点、尘埃颗粒等）图像的互相关，可用来检测和定位这些缺陷。测量经阈值处理得到的图片，可提供关于接触面的数据，以及磨损改变表面后接触面的分布。对空心体积进行骨架化处理，可确定其分配润滑剂的路径的连通性。求出重要的参数后，就可使用这些工具来了解表面行为与历史。

15.13 地形分析和分形维数

上面讨论的传统分析方法的局限性和不足，使得它们不能充分描述真实的表面。它们主要用于机械工程领域的过程控制应用中，此时与先前测量历史的比较可指示出变化，但要求测量技术和仪器不变。由于新设备和新分析方法已出现，因此需要探索它们的完整意义和解释。

人类视觉使用全局地形信息来组织表面上的信息（Scott，1995）。丘陵、山谷、山脊、河道和鞍点的排列方式，包含了用来描述表面的大量信息。地貌或表面可分为由丘陵（从这些点向上的所有路径形成一个特殊的峰）和山谷（从这些点向下的所有路径形成一个凹点）组成的区域。山峰之间的边界称为河道，山谷之间的边界被称为脊线（见图15.43）。

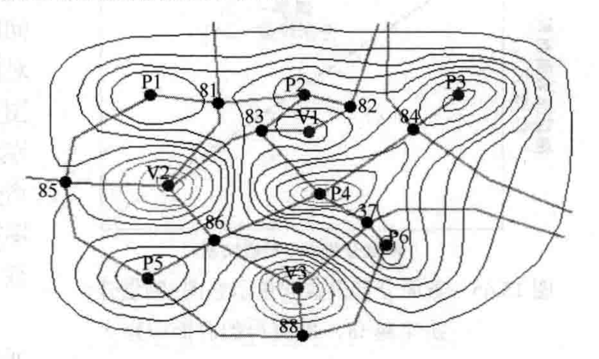


图 15.43　表面等值线图，标记了山峰（红色），山谷（绿色）和鞍点（紫色）

Pfalz 图（Pfalz，1976）或连接山峰和山谷的变更树（见图15.44）总结了地形结构。变更树总结了特征之间的高度差和横向距离，因此可直接删除那些具有垂直和水平小伸展的特征。这是剖面所用图案组合表面的直接扩展。Scott（1997）提出了处理有限伸展真实图像的方法，以及山脊和河道与图像边缘相交时的校正方法。目前，这种信息如何更好地用于表面表征和测量还不清楚，但诸如连通山谷的体积、山谷和山峰在整个表面上的空间分布、河道和山脊的方向这样的参数，可能会很重要。

在很小的局部粗糙度尺度上，许多（但并非全部）表面可由分形维数表示的自相似或自仿射来表征。对其进行测量的方式有几种（但数值上并不完全一致）。必须要有一个参数来描述粗糙度的大小，并要有其他参数来表征表面的方向。分形维数的优点在于，它与测量尺度无关，并会以自然的方式和人类感知粗糙度的方式来汇总表面的“粗糙度”。给定一组表面后，人类解释的表面越粗糙，则分形维数越高。同时，由于分形几何的识别仍是较新的事物，因此人们可能会因“时尚”而不合适地使用它。

在第 11 章中，平面特征的分形维数是沿边界排列的点的函数。维数在 1（线的拓扑维数）和 2（包含特征的平面的拓扑维数）之间。观测长度不随分辨率的增大而变长的平滑边界，其维数为 1.0。形状越粗糙，则其维数越大。

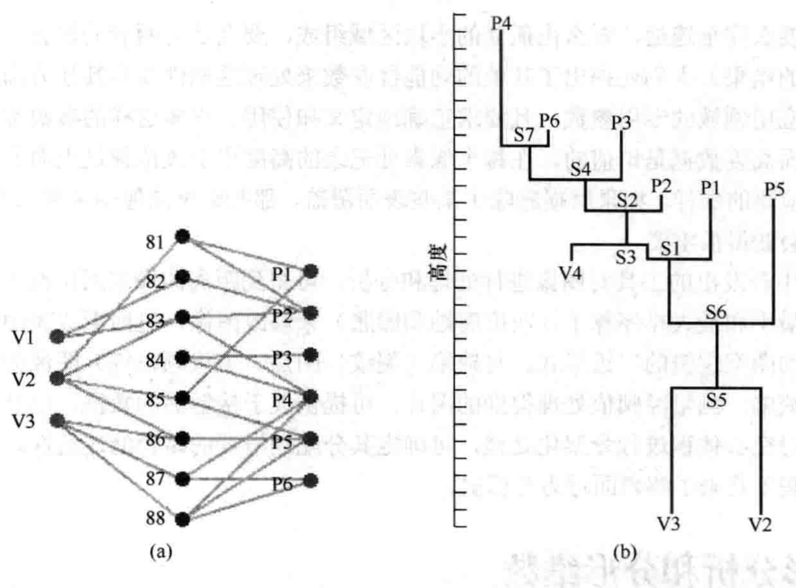


图 15.44 图 15.43 的表面拓扑：(a)Pfalz 图，表明山峰（P）、鞍点（S）和谷地（V）是相邻的；(b)变更树显示了它们之间的高度差和横向距离

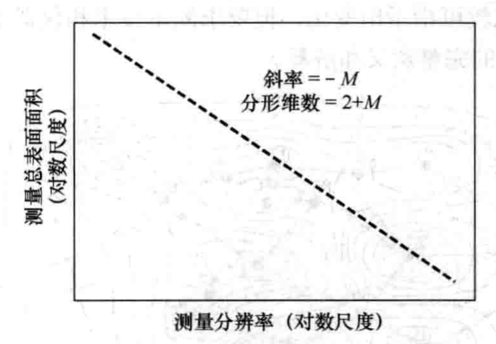


图 15.45 表面分形维数测量示意图：测量分辨率越高，所测得的总面积越大

表面也存在类似的关系。表面的分形维数是一个大于 2（表面的拓扑维数）但小于 3（表面所存在空间的拓扑维数）的实数。非常平滑的表面（维数为 2.0）对应于欧几里得几何，测量区域与测量分辨率的关系图无变化。但对于真实表面而言，放大率增大或分辨率增大时，会显示出更多的细节，同时表面区域也会增大。对于各种自然和人造表面，区域与分辨率关系图在对数图上是线性的，由这条曲线的斜率可求出维数 D，如图 15.45 所示。

这种描述直接来自人们早期关于海岛海岸线、河流长度等的认知，它们的长度都与测量尺度相关。理查得森图图（Mandelbrot，1967）（即英国西海岸长度与测量工具长度的关系图）显示了这种对数关系，它直接促使 Mandelbrot 开始研究自相似结构（在所有尺度上具有相同不规则外观的结构）的数学原理，并将这种行为命名为分形。此后的许多出版物表明，大量的表面具有此类几何特性，且人们研究出了许多方法来测量维度，并开始研究分形维数和表面历史与性能特性间的关系。

以某个范围的分辨率来测量表面区域很难做到；一种方法是通过吸收不同大小分子的单层，然后测量数量。对于许多表面而言，这是不合适的，因为它们不是理想的自相似。对于多数表面而言，水平方向和垂直方向的维数和物理性质是不同的，这意味着不同方向的缩放和自相似性是不同的。在足够大的尺度上，多数表面（甚至地球表面及其山脉）近似为欧几里得形式。对于各向异性的表面，甚至在水平方向上也是不同的。这意味着表面在数学上是自仿射的，而不是自相似的。高度测量是单值的且不能揭示底切这一事实表明，即使是真正自相似的表面，测量的数据也是自仿射的（例如由基底上颗粒有限扩散聚集生成的表面）。对于自仿射表面和数据集，也存在一些正确和有用的测量技术。下面介绍其中的一些实用技术，所有技术的完整介绍见（Russ，1994）。

在多数情形下，即使是对那些非理想分形或非理想各向同性的表面，傅里叶功率谱也可鲁棒地测量其分形维数（Russ，2001）。替代使用功率谱的常用显示模式，幅度与频率对数图表明分形表面是一

条直线，由该直线的斜率可求出维数。使用功率谱图来测量维数的主要缺点是，对于相对平滑的表面，它会过度估计维数的数值（维数为 2.0～2.3），且对于给定尺寸的图像，测量值的精度要低于其他方法的精度。图 15.46 给出了功率谱图的一个例子，由其斜率可得出各向同性表面的分形维数。

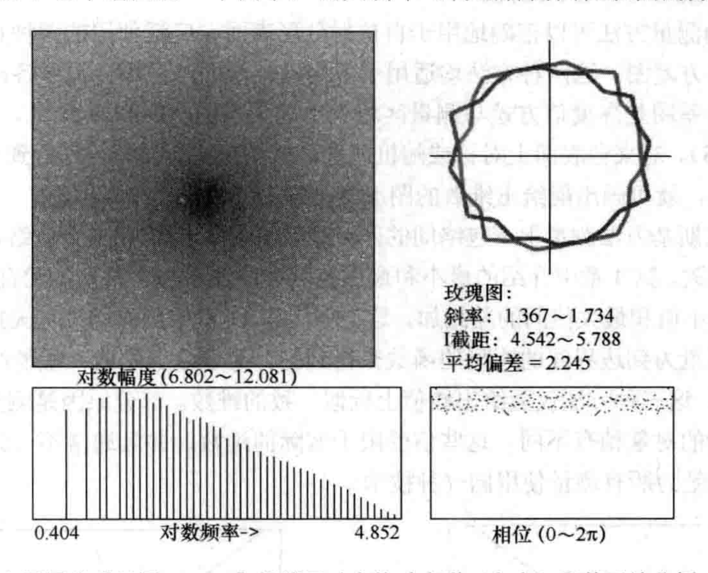


图 15.46　喷丸加工金属图像［见图 15.1(d)］的傅里叶变换功率谱（左上）及其后续分析。在所有方向上平均后的幅度与频率对数关系图（左下）是线性的，这证明由分形行为和直线斜率可得出维数（2.25）。斜率与方向的玫瑰图（右上，红色）表明该表面是各向同性的，在所有方向上的维数相同。玫瑰图上的绿线显示了截距与方向的关系，它是粗糙度的测度，还表明该表面是各向同性的。最后，傅里叶变换中各项的相位分布图（右下）表明它们是均匀随机的，而这正是分形所要求的

生成表面距离图像的二维傅里叶变换，可揭示任何分形的各向异性。这既可以是在各个方向上维数相同但幅度不同的弱各向异性，也可以是维数随方向变化的强各向异性。在幅度与频率对数图中画出斜率和截距与方向的关系图，可定量表示各向异性，如图 15.47 所示。将描述形态的低频数据和揭示仪器限制的高频数据与中间频率数据分开，就可分开表面的分形维数。

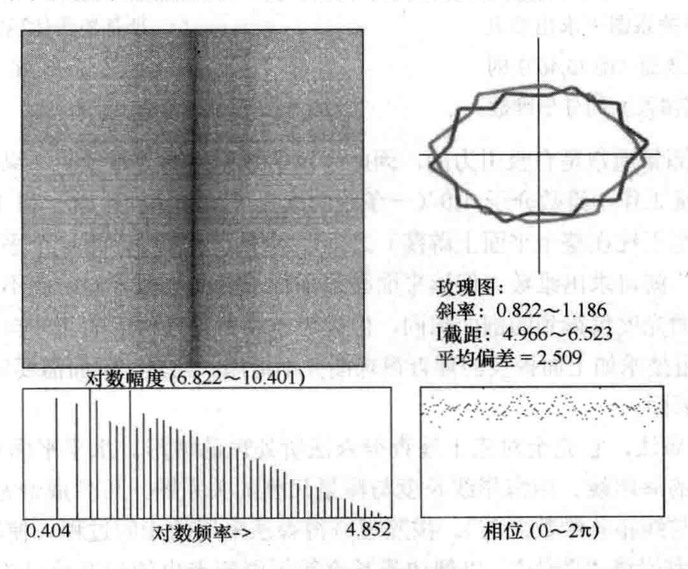


图 15.47　图 15.46 所示的相同数据，但针对的是图 15.1(b)所示的各向异性研磨表面。功率谱图的斜率和截距在水平和垂直方向上是不同的，表明该表面是一个各向异性的分形

就其本身而言，即使是对于理想的分形表面，分形维数也仅是表面粗糙度的部分描述。垂直拉伸表面来增大粗糙度的幅度，不会改变功率谱的斜率或分形维数。需要使用另一个单位长度的测度来表征幅度。功率谱图的截距可以用于此目的。

还有许多其他的测量方法可以正确地用于自仿射分形表面。广泛使用的两种直接处理表面距离图像的技术是盖毯法和方差图。这两种方法均适用于各向同性表面，但不会揭示各向异性，并且不会产生有意义的平均。方差图是高度值方差与测量区域大小的关系图（Dubuc et al.，1989；Russ 1994；Hasegawa et al.，1996）。对来自表面上对称或随机放置区域的值进行平均，可得到一个平均值。以不同的尺寸重复这一处理，就可画出能给出维数的图形（见图 15.48）。

盖毯法或闵可夫斯基方法测量上下包络间的差（对整幅图像求和），上下包络是对表面的拟合，它与所用邻域的大小相关。第 4 章中介绍的最小和最大亮度排序算子与一系列不同直径的邻域结合使用，并对每个直径，将最小值和最大值间的差相加。这类似于第 11 章中剖面的闵可夫斯基维数，它是使用欧氏距离图将面积测量为到边界线的距离的函数来得到的。盖毯法会生成一幅能给出维数的图形，如图 15.49 所示。注意，这三种方法仅能给出数值上近似一致的维数。部分原因是测量的精度有限，部分原因是这些技术测量的对象稍有不同。这些值受限于实际的维数，并且通常不一致，因此在对各个表面进行比较时，应总是为所有测量使用同一种技术。

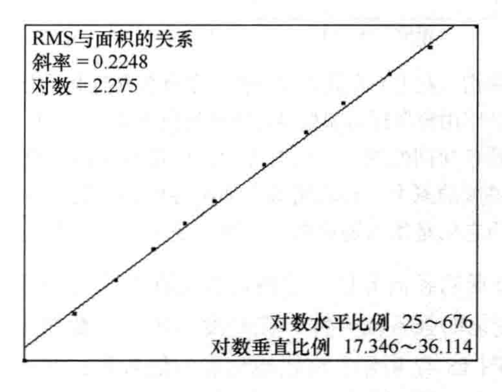


图 15.48 由高度值均方差与邻域大小关系图可求出喷丸加工表面（图 15.46 中的相同图像）的分形维数

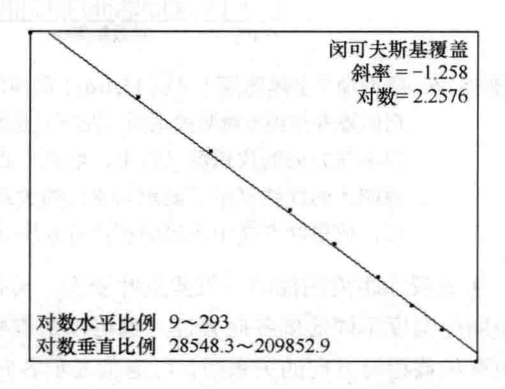


图 15.49 由作为邻域大小函数的闵可夫斯基覆盖体积图可求出分形维数（图 15.46 中的相同图像）

以较低维数进行测量通常是有吸引力的，因此涉及的数据点的数量较少。纵观以往，使用边界线已完成了许多分形测量工作，维数介于 1.0（一条线的欧几里得或拓扑维数）和 1.999…（一条线，其不规则性相当大，以至于线在整个平面上游荡）之间。对于分形表面，用一个平面与该表面相交，然后测量相交线的维数，就可求出维数。但该平面必须平行于名义表面方向，而不是垂直切割。垂直切割生成的剖面与由表面光度仪得到的剖面相同，但因为该表面是自仿射的而非自相似的，因此该剖面的测量很复杂，而常用技术如上面提及的理查得森图并不适用。当然，剖面需要定向到某个特定方向，且不能用于各向异性表面。

水平切割称为切岛法，它完全对应于理查得森法所处理的情形。水平平面对应于海平面，而轮廓是由海面之上岛屿的海岸线。由海岸线长度与测量尺度的关系图，可生成恰好比表面维数小 1 的维数（表面拓扑维数与线拓扑维数之差）。按照理查得森法采用的相同过程（使用圆规卡出某个特殊尺度的长度，并沿着海岸线“跨步”，以便边界长度等于圆规卡出的长度乘以跨步的数量），不可能由计算机中的数字图像测量海岸线的长度，因为点会位于像素地址之间。但许多方法可很容易地在

计算机中实现。

与闵可夫斯基毛毯方法相对应的是闵可夫斯基香肠方法，它将边界线加厚到不同程度，并画出覆盖面积与条带或香肠宽度的关系图。第 11 章中给出的这一方法可非常有效地完成这一任务。另一种方法是计盒法（见图 15.50），即将一个网格放到图像上，将通过边界线的方格数量统计为网格大小的函数。尽管这些方法非常快，但它们仅适用于各向同性表面。绝大多数真实的表面（当然是我们最感兴趣的表面）都不是各向同性的。

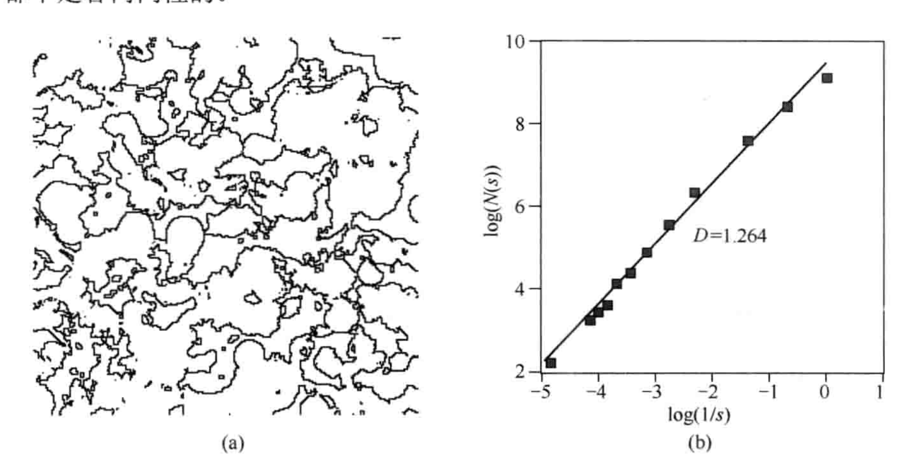


图 15.50　图 15.1(d)所示图像经阈值处理后，在高度范围的中点处生成的等值线，以及计盒测量生成的图形

分形维数和表面各种历史与属性之间的关系，超出了本书的范围。简单地说，涉及脆性破裂或在小区域中沉积大量能量的多数表面成形处理很容易产生分形表面，且维数的数值通常与所涉及的处理相关（Russ，1997）。此外，许多表面接触应用（电子接触、热接触等）取决于接触面积和压力之间的关系，而分形几何与这种情形是相关的。在有些情形下，摩擦和磨损可能也与表面尺寸有关。地球物理侵蚀和沉积过程、云的形成和其他大尺度现象经常由分形几何来表征。表面的分形描述对于表面特征而言是一种新事物。尽管分形描述对某些表面而言是一种强大的方法，但不适用于其他表面，如塑性变形表面。

表面起着非常重要的作用。它们对物体的外观和许多的机械和环境行为负责，包括黏附、摩擦和化学作用。它们既是物体之间相互作用的主要介质，也是人与周围世界互动的主要介质。表面成像与测量采用了大量仪器和技术，而对它们创建的数据集进行分析，可提取出重要的特性参数。未来的目标是综合各种可能性来得到有意义的结果。

参考文献

M. Abbasi-Dezfouli, T. G. Freeman (1994) *Patch matching in stereo-images based on shape* ISPRS International Archives of Photogrammetry and Remote Sensing 30(3/1):1-8.

M. S. Abdaheer, E. Khan (2009) *Shape based classification of breast tumors using fractal analysis* Multimedia Signal Processing and Communication Technologies p. 272-275 IEEE.

G. C. K. Abhayaratne (2006) *Scalable near-lossless image coding* J. Electronic Imaging 15(4):043008.

A. S. Abutaleb (1989) *Automatic thresholding of gray level pictures using two-dimensional entropy* CVGIP 47, 22-32.

S. A. Adhyapak et al. (2007) *Stereo matching via selective multiple windows* J. Electronic Imaging 16(1):013012.

J. Adler et al. (2008) *Replicate-based noise corrected correlation for accurate measurements of colocalization* J. Microscopy 230(1):121-133.

H. Alto et al. (2005) *Content-based retrieval and analysis of mammographic masses* J. Electronic Imaging 14(2):023016.

M. R. Anderberg (1973) Cluster Analysis for Applications Academic Press, New York.

J. R. Anderson et al. (2005) *Segmentation method for three-dimensional visualization of microscopic objects images with a confocal laser scanning microscope* J. Electronic Imaging 15(4):043005.

Agfa Compugraphic Division (1992) Digital Color Prepress (vol. 1 & 2), Agfa Corp., Wilmington MA.

M. Aguilar, E. Anguiano et al. (1992) *Digital filters to restore information from fast scanning tunneling microscopy images* J. Microscopy 165(2):311-324.

E. H. Aigeltinger, K. R. Craig, R. T. DeHoff (1972) *Experimental determination of the topological properties of three dimensional microstructures* J. Microscopy 95:69-81.

S. K. Alibhai, Z. C. Jewell, S. S. Towindo (2001) *The effects of immobilisation on fertility in female black rhino (Diceros bicornis)* J. Zool. 253:333-345.

S. K. Alibhai, Z. C. Jewell (2006) *Individual identification, gender and age class determination for the lowland tapir (Tapirus terrestris) using a footprint identification technique (FIT)* Report to P. Medici, Chair, IUCN/SSC Tapir Specialist Group.

S. K. Alibhai, Z. C. Jewell (2007) *A census estimate of lowland tapir (Tapirus terrestris) in the Morro de Diabo State Park, Sao Paulo State, Brazil, using the Footprint Identification Technique (FIT)* Report to P. Medici, Chair, IUCN/SSC Tapir Specialist Group.

S. K. Alibhai, Z. C. Jewell (2008) *Identifying white rhino (Ceratotherium simum) by a footprint identification technique, at the individual and species levels* Endangered Species Research 4:219-225.

Y. Anguy et al. (2005) *Practical modeling of porous media from second-order statistics: the phase-retrieval problem and the interpretation of covariance* Microscopy 220(3):140-149.

G. R. Arce, J. L. Paredes, J. Mullan (2000) *Nonlinear flitering for image analysis and enhancement* in Handbook of Image and Video Processing (A. Bovic, ed.), Academic Press, San Diego.

J. Arvo (ed.) (1991) Graphics Gems II Academic Press, San Diego, CA.

J. Astola, P. Haavisto, Y. Neuvo (1990) *Vector median filters* Proc IEEE 78:678-689.

A. Averbuch, Y. Keller (2002) *FFT based image registration* IEEE Int. Conf. Acoustics, Speech, and Signal Processing 4:3601-3608.

N. Baba et al. (1984) *Computer-aided stereographic representation of an object reconstructed from micrographs of serial thin sections* Micron and Microscopica Acta 15:221-226.

N. Baba et al. (1989) *Serial section reconstruction using a computer graphics system: Application to intracellular structures in yeast cells and to the periodontal structure of dogs'teeth* Journal of Electron Microscopy Technique 11:16-26.

J. Bacus et al. (1995) *Method and apparatus for automated assay of biological specimens* U. S. Patent 5. 428, 690.

J. V. Bacus, J. W. Bacus (2000) *Method and apparatus for acquiring and reconstructing magnified specimen images from a computer-controlled microscope* U. S. Patents 6, 101, 265; 6, 626, 392; 6, 272, 235; 6, 404, 906; 6, 522, 774; 6, 775, 402.

J. W. Bacus, J. V. Bacus (2002) *Method and apparatus for internet, intranet and local viewing of virtual microscope slides* U. S. Patents 6, 396, 941; 6, 674, 881.

A. J. Baddeley, H. J. G. Gundersen, L. M. Cruz-Orive (1986) *Estimation of surface area from vertical sections* Microscopy 142:259-276.

A. J. Baddeley et al. (1987) *Three-dimensional analysis of the spatial distribution of particles using the tandemscanning reflected light microscope* Acta Stereol. 6(suppl II): 87-100.

X. Bai et al. (2008) *New class of top-hat transformation to enhance infrared small targets* J. Electronic Imaging Letters 17(3):030501.

J. W. Baish, R. K. Jain (2000) *Fractals and cancer*, Cancer Res 61(22):8347-8350.

Y. Balagurunathan, E. R. Dougherty (2003) *Morphological quantification of surface roughness* Optical Engineering 42(6):1795-1804.

R. Balasubramanian, C. A. Bouman, J. P. Allebach (1994) *Sequential scalar quantization of color images* Journal of Electronic Imaging 3(1):45-59.

D. Ballard (1981) *Generalizing the Hough transform to detect arbitrary shapes* Pattern Recognition 13(2): 111-122.

D. H. Ballard, C. M. Brown (1982) Computer Vision, Prentice Hall, Englewood Cliffs, NJ.

D. H. Ballard, C. M. Brown, J. A. Feldman (1984) *An approach to knowledge-directed image analysis* in Computer Vision Systems (A. R. Hanson, E. M. Riseman, ed.) Academic Press, New York, 271-282.

S. T. Barnard, W. B. Thompson (1980) *Disparity analysis of images* IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-2(4):333-340.

F. L. Barnes et al. (1990) *Geometric Effects in Tomographic Reconstruction* Lawrence Livermore National Laboratory Rept UCRL-ID-105130.

M. F. Barnsley(1988) Fractals Everywhere, Academic Press, Boston.

M. F. Barnsley et al. (1986) *Solutions of an inverse problem for fractals and other set* Proc. Nat'l Acad. Sci. 83:1975-1977.

M. F. Barnsley, A. D. Sloan (1991) *Method and apparatus for processing digital data* U. S. Patent 5, 065, 447.

P. G. J. Barten (1999) Contrast Sensitivity of the Human Eye and its Effects on Image Quality SPIE Press, Bellingham, WA.

E. Barth et al. (1998) *Intrinsic 2d features as textons* J. Optical Society of America A157:1723-1732.

H. G. Barth, S. -T. Sun (1985) *Particle size analysis* Anal. Chem. 57:151.

M. Barth et al. (1989) *Approximation of missing-cone data in 3D electron microscopy* Ultramicroscopy 31:365.

S. Bartling et al. (2005) *Registration and fusion of CT and MRI of the temporal bone* J. Comput. Assist. Tomography 29(3):305-310.

A. Barty et al. (2000) *Quantitative phase tomography*, Optics Communications 175:329-336.

S. Battiato et al. (2003) *High dynamic range imaging for digital and still camera: an overview* J. Elect. Imaging 12(3):459-469.

W. Bauer, C. D. Mackenzie (1995) *Cancer detection via determination of fractal cell dimension*, Michigan State Univ. report CL-980.

S. Bayram et al. (2006) *Image manipulation detection* J. Electronic Imaging 15(4):041102.

G. Bearman (2003) *Biological imaging spectroscopy* in Biomedical Photonics Handbook (T. V0-Dinh, ed.) CRC Press, Boca Raton, FL.

A. L. D. Beckers, A. W. M. Smeulders (1989) *The probability of a random straight line in two and three dimensions* Pattern Recognition Letters 11(4):233-240.

J. K. Beddow, G. C. Philip, A. F. Vetter (1977) *On relating some particle profiles characteristics to the profile Fourier coefficients* Powder Technol. 18:15-19.

M. Beil et al. (2005) *Quantitative analysis of keratin filament networks in scanning electron microscopy images of cancer cells* J. Microscopy 220(2):84-95.

P. N. Belhumeur et al. (2008) *Searching the World's Herbaria: A System for Visual Identification of Plant Species*, in D. Forsyth et al. (eds.) ECCV 2008 Part IV, LNCS 5305, p. 116-119, Springer Verlag, Berlin.

A. Bennia, S. M. Riad (1992) *Filtering capabilities and convergence of the Van-Cittert deconvolution technique* IEEE Transactions on Instrumentation and Measurement 41(2):246-250.

R. Bergman et al. (2008) *Comprehensive solutions for automatic removal of dust and scratches from images* J. Electronic Imaging 17(1):013010.

R. Bergman et al. (2008) *Detection of textured areas in natural images using an indicator based on component counts* J. Electronic Imaging 17(4):043003.

J. G. Berryman, S. C. Blair (1986) *Use of digital image analysis to estimate fluid permeability of porous materials: Application of two-point correlation function* J. App. Phys. 60(6):1930-1938.

M. Bertero and P. Boccacci (1998) Introduction to Inverse Problems in Imaging Institute of Physics Publishing.

G. Bertrand, J. -C. Everat, M. Couprie (1997) *Image segmentation through operators based on topology* J. Electronic Imaging 6(4):395-405.

V. Berzins (1984) *Accuracy of Laplacian edge detectors* Comput. Vis. Graph. Image Proc. 27:1955-2010.

P. J. Besl, N. D. McKay (1992) *A method for registration of 3-D shapes*, IEEE Trans. Patt. Anal. Mach. Intell. 14(2): 239-256.

O. Betz et al. (2007) *Imaging applications of synchrotron X-ray phase-contrast microtomography in biological morphology and biomaterials science* J. Microscopy 227(1):51-71.

S. Beucher, C. Lantejoul (1979) *Use of Watersheds in Contour Detection* Proc. Int'l. Workshop on Image Proc., CCETT, Rennes, France.

V. Bhaskaran, K. Konstantinides (1997) Image and Video Compression Standards: Algorithms and Architectures Springer, New York.

G. B. Binnig, C. F. Quate, C. Gerber (1986) *Atomic force microscope*, Phys. Rev. Lettr. 52:930.

J. L. Bird, D. T. Eppler, D. M. Checkley, Jr. (1986) *Comparisons of herring otoliths using Fourier series shape analysis* Can. J. Fish. Aquat. Sci. 43:1228-1234.

K. S. Birdi (2003) Scanning Probe Microscopies CRC Press, Boca Raton, FL.

C. M. Bishop (1996) Neural Networks for Pattern Recognition Oxford Univ. Press, Oxford, UK.

C. M. Bishop (2007) Pattern Recognition and Machine Learning Springer, New York.

W. J. Black (1986) Intelligent Knowledge-based Systems: An Introduction Van Nostrand Reinhold, London.

R. A. S. Blackie et al. (1987) *Reconstruction of three dimensional images of microscopic objects using holography* Proc. Royal Microscopical Society 22:98.

F. Boddeke (1998) *Quantitative Fluorescence Microscopy*, ASCI Dissertation Series, Delft Univ. Press, Delft Netherlands.

F. R. Boddeke et al. (1994) *Autofocusing in microscopy based on the OTF and sampling* Bioimaging 2:193-203.

J. Bollmann et al. (2004) *Automated particle analysis: calcareous microfossils* p. 229-252 in E. Francus (ed.), Image Analysis, Sediments and Paleoenvironments Springer, Dordrecht, Netherlands.

J. S. de Bonet, P. Viola (1998) *Texture recognition using a nonparametric multi-scale statistical model* CVPR '98: Proc. IEEE Computer Soc. Conf. Computer Vision Patt. Recog., p. 641.

N. Bonnet et al. (1992) *New applications of multivariate statistical analysis in spectroscopy and microscopy* Ultramicroscopy 40:1.

M. J. Booth et al. (2010) *Full spectrum filterless fluorescence microscopy*, J. Microscopy 237(1):103-109.

G. Borgefors (1996) *On digital distance transforms in three dimensions* Computer Vision and Image Understanding 64(3):368-376.

G. Borgefors, I. Nystron, G. Sanniti di Baja (1999) *Computing skeletons in three dimensions*, Pattern Recognition 32(7): 1225-1236.

S. -T. Bow (1992) Pattern Recognition and Image Preprocessing Marcel Dekker, New York.

E. T. Bowman et al. (2000) *Particle shape characterisation using Fourier analysis* Geotechnique 51(6):545-554.

A. Boyde (1973) *Quantitative photogrammetric analysis and quantitative stereoscopic analysis of SEM images* J. Microscopy 98:452.

A. Boyde (2004) *Improved depth of field in the scanning electron microscope derived from through-focus image stacks* Scanning 26:265-269.

B. Bracegirdle, S. Bradbury (1995) Modern Photomicrography Bios Scientific Publishers, Oxford, UK.

R. N. Bracewell (1984) *The fast Hartley transform* Proc IEEE 72:8.

R. N. Bracewell (1986) The Hartley Transform Oxford Univ. Press, Oxford, UK.

R. N. Bracewell (June 1989) *The Fourier transform* Scientific American.

S. Bradbury, B. Bracegirdle (1998) Introduction to Light Microscopy Bios Scientific Publishers, Oxford, UK.

M. von Bradke et al. (2004) *Porosity determination of ceramic materials by digital image analysis — a critical examination* Scanning 27:132-135.

G. Braudaway (1987) *A procedure for optimum choice of a small number of colors from a large color palette for color imaging* Proc. Electronic Imaging '87, San Francisco, CA.

S. J. Brazdilova, M. Kozubek (2009) *Information content analysis in automated microscopy imaging using an adaptive autofocus algorithm for multimodal functions* J. Microscopy 236(2):194-202.

J. E. Bresenham (1965) *Algorithm for computer control of a digital plotter* IBM Systems Journal 4(1):25-30.

J. E. Bresenham (1977) *A linear algorithm for incremental digital display of circular arcs* Communications of the ACM 20(2).

L. G. Briarty, P. H. Jenkins (1984) *GRIDSS: an integrated suite of microcomputer programs for three-dimensional graphical reconstruction from serial sections* J. Microscopy 134:121-124.

D. S. Bright, E. B. Steel (1986) *Bright field image correction with various image processing tools in* Microbeam Analysis 1986 (A. D. Romig, ed.) San Francisco Press, p. 517-520.

D. S. Bright, D. E. Newbury, R. B. Marinenko (1988) *Concentration-concentration histograms: scatter diagrams applied to quantitative compositional maps* in Microbeam Analysis 1988 (D. E. Newbury ed.) San Francisco Press, p. 18-24.

A. D. Brink, N. E. Pendcock (1989) *Minimum cross-entropy threshold selection* Pattern Recognition 29, 179-188.

P. Brodatz (1966) Textures: a Photographic Albun for Artists and Designers Dover, New York.

L. G. Brown (1992) *A Survey of image registration techniques* ACM Computing Surveys 24(4):325-376.

M. Brown, D. G. Lowe (2003) *Recognizing panoramas* International Conference on Computer Vision, Nice, France, pp 1218-1225.

M. Brown et al. (2005) *Multi-image matching using multi-scale oriented patches* International Conference on Computer Vision and Pattern Recognition San Diego, CA.

S. F. Brown (2007) *Seeing triple* Scientific American 296(6):86-89.

D. Bruton (2005) <www. physics. sfasu. edu/astro/color. html> (web page).

R. K. Bryan, J. Skilling (1980) *Deconvolution by maximum entropy, as illustrated by application to the jet of M87*, Mon. Not. R. Ast. Soc. 191:69-79.

M. A. Buena-Ibarra et al. (2005) *Fast autofocus algorithm for automated microscopes* Optical Engineering 44(6):063601.

M. Bueno et al. (2005) *Polychromatic image fusion algorithm and fusion metric for automatized microscopes* Optical Engineering 44(9):093201.

V. Buzuloiu, M. Ciuc, R. M. Rangayyan, C. Vertan (2001) *Adaptive-neighborhood histogram equalization of color images* Journal of Electronic Imaging 10(2):445-459.

P. Campisi, K. Eglazarian (2007) Blind Image Deconvolution: Theory and Applications CRC Press, Boca Raton, FL.

J. Canny (1986) *A computational approach to edge detection* IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) 8(6):679-698.

A. S. Carasso (2001) *Direct blind deconvolution* SIAM J. Appl. Math. 61:1980-2007.

A. S. Carasso (2002) *The apex method in image sharpening and the use of low exponent Levy stable laws* SIAM J. Appl. Math. 63(2):598-618.

A. S. Carasso (2006) *APEX blind deconvolution of color Hubble space telescope imagery and other astronomical data* Optical Engineering 45(10):107004.

I. C. Carlsen (1985) *Reconstruction of true surface topographies in scanning electron microscopes using backscattered electrons* Scanning 7:169-177.

K. Carlsson (1990) *Scanning and detection techniques used in a confocal scanning laser microscope* J. Microscopy 157:21-27.

W. A. Carrington (1990) *Image restoration in 3D microscopy with limited data* in Bioimaging and Two Dimensional Spectroscopy, Proc. SPIE vol. 1205 (L. C. Smith, ed.), 72-83.

J. E. Castle, P. A. Zhdan (1997) *Characterization of surface topography by SEM and SFM: Problems and solutions* J. Phys. D: App. Phys. 30:722.

J. E. Castle et al. (1998) *Enhanced morphological reconstruction of SPM images* J. Phys. D: Appl. Phys 31: 3437-3445.

K. R. Castleman (1979) Digital Image Processing Prentice Hall, Englewood Cliffs, NJ.

K. R. Castleman (1996) Digital Image Processing, Prentice Hall, Upper Saddle River, NJ.

R. G. Caves et al. (1992) *Matching map features to synthetic aperture radar (SAR) images using template matching* IEEE Trans. Geosci. Remote Sens. 30(4):680-685.

R. L. T. Cederberg (1979) *Chain-link coding and segmentation for raster scan devices* Comput. Graph. Image Proc. 10:224-234.

D. Chappard et al. (2005) *Microcomputed tomography for the study of hard tissues and bone biomaterials* Microscopy and Analysis 5:23-25.

M. E. Celebi et al. (2007) *Nonlinear vector filtering for impulsive noise removal from color images* J. Electronic Imaging 16(3):033008.

M. E. Celebi, Y. A. Aslondogan (2008) *Robust switching vector median filter for impulsive noise removal* J. Electronic Imaging 17(4):043006.

P. -K. Chan et al. (2007) *Automated segmentation in confocal images using a density clustering method* J. Electronic Imaging 16(4):043003.

Y. -S. Chen et al. (2005) *Efficient fuzzy c-means clustering for image data* J. Electronic Imaging 14:013017.

Y. -S. Chen et al. (2008) *Image superresolution using fractal coding* Optical Engineering 47(1):017007.

P. Chieco et al. (1994) *A user's guide for avoiding errors in absorbance image cytometry* Histochemical Journal 26:1-19.

M. S. Choi, W. Y. Kim (2002) *A novel two stage template matching method for rotation and illumination invariance* Pattern Recogn. 35(1):119-129.

S. K. Chow et al. (2006) *Automated microscopy system for mosaic acquisition and processing* J. Microscopy 222(2):76-84.

C. Christopoulos, A. Skodras, and T. Ebrahimi (2000) *The JPEG2000 still image coding system: An overview* IEEE Trans. Consum. Electron. 46(4):1103-1127.

C. K. Chui (1992) An Introduction to Wavelets Academic Press, London.

P. H. van Cittert (1931) *Effect of slit width upon distribution of intensity in spectral lines, Part II Z.* Phys. 69:298-308.

F. Ciurea, B. Funt (2004) *Tuning Retinex parameters* J. Electronic Imaging 13(1):58-64.

J. R. B. Cockett (1987) *Decision expression optimization* Fundamenta Informaticae X, 93-114.

L. Cohen (1991) *On active contour models and balloons* Computer Vision, Graphics and Image Processing 53(2):211-218.

M. L. Comer, E. J. Delp (1999) *Morphological operations for color image processing*, J. Electronic Imaging 8:279-289.

J. Condeco et al. (2000) *A comparative study of image stitching algorithms for surface topography measurements* X. Int'l Colloq. on Surfaces, Germany.

J. Cookson (1994) *Three-dimensional reconstruction in microscopy* Proceedings of the Royal Microscopical Society, 29(1) Jan., 1994, p. 3-10.

J. W. Cooley, J. W. Tukey (1965) *An algorithm for the machine calculation of complex Fourier series* Mathematics of Computation.

T. J. Cooper, F. A. Baqal (2004) *Analysis and extensions of the Frankle-McCann Retinex algorithm* J. Electronic Imaging 13(1):85-92.

A. M. Cormack (1963) *Representation of a function by its line integrals with some radiological applications* J. Appl. Phys. 34:2722-2727.

A. M. Cormack (1964) *Representation of a function by its line integrals with some radiological applications II* J. Appl. Phys. 35:2908-2913.

T. N. Cornsweet (1970) Visual Perception Academic Press, New York.

L. F. Costa, R. M. Cesar (2001) Shape Analysis and Classification, CRC Press, Boca Raton, FL.

L. Costaridou (ed.) (2005) Medical Image Analysis Methods, CRC Press, Boca Raton, FL.

M. Coster, J. -L. Chermant (1985) Précis D'Analse D'Images Éditions du Centre National de la Recherche Scientifique, Paris.

G. Cox (2007) Optical Imaging Techniques in Cell Biology CRC Press, Boca Raton, FL.

H. D. Crane, T. P. Piantanida (1983) *On seeing reddish green and yellowish blue* Science 221:1078-1080.

L. -M. Cruz-Orive (1976) *Particle size-shape distributions: the general spheroid problem* J. Microscopy 107:235 and 112:153.

L. -M. Cruz-Orive (1983) *Distribution-free estimation of sphere size distributions from slabs showing overprojections and truncation, with a review of previous methods* J. Microscopy 131:265.

L. -M. Cruz-Orive (1993) *Systematic sampling in stereology* Proc. 49th Session, International Statistical Institute (Book 2), 451-468.

L. -M. Cruz-Orive (2005) *A new stereological principle for test lines in three-dimensional space* J. Microscopy 219(1):18-28.

P. E. Danielsson (1980) *Euclidean distance mapping* Comput. Graph. Image Proc. 14:227-248.

H. J. Dartnall et al. (1983) *Human visual pigments: microspectrophotometric results from the eyes of seven persons* Proceedings of the Royal Society of London, B 220:115-130.

R. Datta, D. Joshi, J. Li, J. Z. Wang (2008) *Image retrieval: Ideas, influences and trends of the new age* ACM Computing Surveys 40(2):1-60.

I. Daubechies (1992) *Ten Lectures on Wavelets* CBMS-NSF Regional Conf. Series Applied Math, Philadelphia.

I. Daubechies (1996) *Where do wavelets come from? — A personal point of view* Proc. IEEE 84(4):510-513.

D. G. Daut et al. (1993) *Double predictor differential pulse coded modulation algorithm for image data compression* Optical Engineering 32(7):1514-1523.

J. Davidson (1991) *Thinning and skeletonization: a tutorial and overview* in Digital Image Processing: Fundamentals and Applications (E. Dougherty, ed.) Marcel Dekker, New York.

E. R. Davies (1988) *On the noise suppression and image enhancement characteristics of the median, truncated median and mode filters* Pattern Recognition Letters 7:87-97.

P. E. Debevec, J. Malik (1997) *Recovering high dynamic range radiance maps from photographs* SIGGRAPH 1997.

H. W. Deckman et al. (1989) *Microtomography detector design* Advances in X-ray Analysis 32:641.

R. T. Dehoff, F. N. Rhines (1968) Quantitative Microscopy McGraw Hill, New York.

D. DeMandolx, J. Davoust (1997) *Multicolor analysis and local image correlation in confocal microscopy* J. Microscopy 185:21-36.

D. R. Denley (1990a) *Scanning tunneling microscopy of rough surfaces* J. Vac. Sci. Techno. A8(1):603-607.

D. R. Denley (1990b) *Practical application of scanning tunneling microscopy* Ultramicroscopy 33:83-92.

S. D. Deshpande et al. (1989) *Max-mean and max-median filters for detection of small-targets* Proc. SPIE 3809, 74-83.

P. A. Devijver, J. Kittler (1980) Pattern Recognition: A Statistical Approach Prentice-Hall, Englewood Cliffs, NJ.

N. Dey et al. (2006) *Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution* Microscopy Research and Technique 69:260-266.

G. Diaz et al. (1989) *Elliptic Fourier analysis of cell and nuclear shapes* Comput. Biomed. Res. 22:405-414.

G. Diaz et al. (1990) *Recognition of cell surface modulation by elliptic Fourier analysis* Computer Methods and Programs in Biomedicine 31:57-62.

M. Dietzsch et al. (1997) *The MOTIF method (ISO 12085) — a suitable description for functional manufactural and*

metrological requirements in 7th Int'l Conf. on Metrology and Properties of Engineering Surfaces (B. G. Rosen, R. J. Crafoord, eds.) Chalmers Univ., Göteborg Sweden, pp. 231-238.

E. R. Dougherty, J. Astola (1994) An Introduction to Nonlinear Image Processing SPIE, Bellingham, WA.

E. R. Dougherty, J. Astola (1999) Nonlinear Filters for Image Processing, SPIE, Bellingham, WA.

N. Draper, H. Smith (1981) Applied Regression Analysis (2nd edition) Wiley, New York.

A. Drygajlo (1992) *Butterfly orthogonal structure for fast transforms, filter banks and wavelets* IEEE Conf. Acoustics, Speech, and Signal Processing, 5:81-84.

B. Dubuc et al. (1989) *Evaluating the fractal dimension of profiles* Phys. Rev A 39:1500-1512.

R. O. Duda, P. E. Hart (1972) *Use of the Hough transform to detect lines and curves in pictures* Communications of the ACM 15(1):11-15.

R. O. Duda, P. E. Hart (1973) Pattern Classification and Scene Analysis Wiley, New York.

F. Durand, J. Dorsey (2002) *Fast bilateral filtering for the display of high dynamic range images* ACM Trans. Graphics 2002.

T. R. Edwards (1982) *Two-dimensional convolute integers for analytical instrumentation* Anal. Chem. 54:1519-1524.

M. Egmont-Petersen et al. (2002) *Image processing with neural networks—a review* Pattern Recognition, 35(10): 2279-2301.

P. A. van den Elsen et al. (1995) *Automatic Registration of CT and MR brain images using correlation of geometrical features* IEEE Transactions on Medical Imaging 14(2):384-396.

P. A. van den Elsen et al. (1994) *Grey value correlation techniques used for automatic matching of CT and MR brain and spine images*, SPIE Int'l Conf. on Visualization in Biomedical Computing 2359:227-237.

P. A. van den Elsen et al. (1993) *Medical image matching — A review with classication*, IEEE Engineering in Medicine and Biology 12(1):26-39.

R. Ehrlich, B. Weinberg (1970) *An exact method for characterization of grain shape* J. Sediment. Petrol. 40:205-212.

R. Ehrlich et al. (1984) *Petrographic image analysis: 1. Analysis of reservoir pore complexes* J. Sed. Petrol. 54:1365-1378.

T. Ehtezazi et al. (1999) *Determination of the internal morphology of poly (D, L-lactide) microspheres using stereological methods* Journal of Controlled Release 57:301-314.

M. Emmenlauer et al. (2009) *Xuv tools: free, fast and reliable stitching of large 3D datasets* J. Microscopy 233(1):42-60.

A. Engel, A. Massalski (1984) *3D reconstruction from electron micrographs: Its potential and practical limitations* Ultramicroscopy 13:71-84.

R. Estrada, C. Tomasi (2009) *Manuscript bleed-through removal via hysteresis thresholding* Proc. 10th Int'l Conf. on Document Analysis and Recognition, IEEE, 753-757.

R. Fabbri et al. (2008) *2D Euclidean distance transform algorithms: A comparative survey* ACM Computing Surveys 40(1):2:1-44.

Y. Fahmy et al. (1991) *Application of fractal geometry measurements to the evaluation of fracture toughness of brittle intermetallics* J. Mater. Res. 6(9):1856-1861.

C. Faloutsos et al. (1994) *Efficient and effective querying by image content* Journal of Intelligent Information Systems, 3:231-262.

H. Farid (2008) *Digital image forensics* Scientific American 298(6):66-71.

R. Fattal (2007) *Image upsampling via imposed edge* statistics ACM Trans. Graphics (Proc. SIGGRAPH 2007), 26(3):96-102.

J. Feder (1988) Fractals Plenum Press, New York.

L. A. Feldkamp et al. (1984) *Practical cone beam algorithm* J. Opt. Soc. Am. 1(6):612.

S. F. Ferson et al. (1985) *Measuring shape variation of two-dimensional outlines* Systematic Zoology 34: 59-68.

J. Fiala (2005) *Reconstruct: a free editor for serial section microscopy* J. Microscopy 218(1):52-61.

L. F. Firestone, K. Cook et al. (1991) *Comparison of autofocus methods for automated microscopy* Cytometry 12:195-206.

Y. Fisher et al. (1992) *Fractal image compression using iterated transforms* in James A. Storer, ed. Image and Text Compression, pages 35-61. Kluwer Academic Publishers, Boston, MA.

M. Flickner et al. (1995) *Query by image and video content: the QBIC system*, IEEE Computer 28(9):23-32.

A. G. Flook (1978) *Use of dilation logic on the Quantimet to achieve fractal dimension characterization of texture and structured profiles* Powder Techn. 21:295-298.

A. G. Flook (1982) *Fourier analysis of particle shape* in Particle Size Analysis 1981-2 (N. G. Stanley-Wood, T. Allen, ed.) Wiley Heyden, London.

J. D. Foley, A. Van Dam (1984) Fundamentals of Interactive Computer Graphics Addison Wesley, Reading, MA.

J. D. Foley, A. Van Dam (1995) Fundamentals of Interactive Computer Graphics, 2nd Edition Addison Wesley, Reading, MA.

B. Forster et al. (2004) *Complex wavelets for extended depth-of-field: A new method for the fusion of multichannel microscopy images* Microscopy Research and Technique 65(1-2):33-42.

I. France et al. (2004) *Software aspects of automated recognition of particles: the example of pollen* pp 253-272 in E. Francus (ed.) Image Analysis, Sediments and Paleoenvironments Springer, Dordrecht, Netherlands.

E. Francus (ed.) (2004) Image Analysis, Sediments and Paleoenvironments Volume 7, Springer, Dordrecht, Netherlands.

P. Francus, E. Pinard (2004) *Testing for sources of errors in quantitative image analysis* pp 87-104 in E. Francus (ed.) Image Analysis, Sediments and Paleoenvironments Volume 7, Springer, Dordrecht, Netherlands.

J. Frank, ed. (1992) Electron Tomography, Plenum Press.

R. J. Frank et al. (1995) *Voxelwise percentage tissue segmentation of human brain magnetic resonance images* in Abstracts, 25th Annual meeting, Soc. for Neuroscience, Washington, 694, cited in M. Sonka et al. (1999) Image Processing, Analysis and Machine Vision Brooks Cole, Pacific Grove, CA, p. 132.

M. Frederik et al. (1997) *Multi-modality image registration by maximization of mutual information* IEEE Transactions on Medical Imaging 16(2):187-198.

H. Freeman (1961) *On the encoding of arbitrary geometric configurations* IEEE Trans. EC-10:260-268.

H. Freeman (1974) *Computer processing of line-drawing images* Comput. Surveys 6:57-97.

H. Freeman, L. S. Davis (1977) *A corner finding algorithm for chain-code curves* IEEE Trans. Comput. 26: 297-303.

W. T. Freeman et al. (2001) *Example-based super-resolution*, IEEE Computer Graphics and Applications, 22(2):56-65.

T. Freeth (2009) *Decoding an ancient computer*, Scientific American 301(6):76-83.

W. Frei (1977) *Image enhancement by histogram hyperbolization* CGIP 6(3):286-294.

W. Frei, C. C. Chen (1977) *Fast boundary detection: A generalization and a new algorithm* IEEE Trans. Comput. C-26:988-998.

B. Freisleben, A. Schrader (1997) *Color quantization with a hybrid genetic algorithm* Proc. 6th IEEE Conf. On Image Proc., Dublin, 89-93.

B. R. Frieden (1988) *A comparison of maximum entropy, maximum a postiori and median window restoration algorithms*, in Scanning Microscopy Suppl. 2 (P. Hawkes et al., ed.), pp. 107-111, Scanning Microscopy International, Chicago.

M. Frigo, S. G. Johnson (2005) *The design and implementation of FFTW3* Proceedings of the IEEE 93(2): 216-231.

J. P. Frisby (1980) Vision: Illusion, Brain and Mind Oxford Univ. Press, Oxford, U. K.

J. P. Frisby and J. V. Stone (2010) Seeing, 2nd Edition: The Computational Approach to Biological Vision MIT Press, Boston, MA.

K. S. Fu (1974) Syntactic Methods in Pattern Recognition Academic Press, Boston.

K. S. Fu, J. K. Mui (1981) *A survey on image segmentation* Pattern Recognition 13:3-16.

K. S. Fu(傅京孙) (1982) Syntactic Pattern Recognition and Applications Prentice-Hall, Englewood Cliffs, NJ.

H. Fuchs et al. (1982) *Adding a true 3-D display to a raster graphics system* IEEE Comput. Graphics Applic. 2:73-78.

K. Fukunaga (1990) Statistical Pattern Recognition 2nd Edition, Academic Press, Boston.

B. Funt et al. (2004) *Retinex in MATLAB* J. Elect. Imaging 13(1):48-57.

J. E. Gardi et al. (2008) *Automatic sampling for unbiased and efficient stereological estimation using the proportionator in biological studies* J. Microscopy 230(1):108-120.

R. S. Gentile et al. (1990) *Quantization of color images based on uniform color spaces* J. Imaging Technology 16(1):12-21.

S. Geuna (2005) *The revolution of counting "tops": two decades of the Disector principle in morphological research* Microscopy Research and Technique 66:270-274.

O. Ghita et al. (2005) *Computational approach for depth from defocus* J. Electronic Imaging 14(2):023021.

J. Glaser et al. (2007) Stereology for Biological Research with a Focus on Neuroscience, MBF Press, Burlington, VA.

A. M. Gokhale et al. (2004) *Design-based estimation of surface area in thick tissue sections of arbitrary orientation using virtual cycloids* Microscopy 216(1):25-31.

T. H. Goldsmith (2006) *What birds see* Scientific American 294(6):68-75.

O. Gomes, S. Paciornik (2005) *Automatic classification of graphite in cast iron* Microscopy and Microanalysis 11:363-371.

R. C. Gonzalez, P. Wintz (1987) Digital Image Processing, 2nd Edition, Addison Wesley, Reading, MA.

R. C. Gonzalez, R. E. Woods (1993) Digital Image Processing, Addison Wesley, Reading, MA.

R. C. Gonzalez et al. (2004) Digital Image Processing Using Matlab Prentice Hall, Upper Saddle River, NJ.

M. F. Goodchild (1980) *Fractals and the accuracy of geographical measures* Mathematical Geology 12(2): 85-98.

R. Gordon (1974) *A tutorial on ART (Algebraic Reconstruction Techniques)* IEEE Trans NS-21:78-93.

A. Goshtasby (2005) 2-D and 3-D Image Registration for Medical, Remote Sensing and Industrial Applications Wiley, Hoboken, NJ.

A. C. Gould et al. (1990) *The atomic force microscope: a tool for science and industry* Ultramicroscopy 33:93.

Y. A. Gowayed, J. C. Russ (1991) *Geometric characterization of textile composite preforms using image analysis techniques*, Journal of Computer Assisted Microscopy 3(4):189-200.

G. H. Granlund (1972) *Fourier preprocessing for hand print character recognition* IEEE Trans. Comput. C21(2):195-201.

A. Grasselli (ed.) (1969) Automatic Interpretation and Classification of Images Academic Press, Boston.

C. Gratin, F. Meyer (1992) *Morphological three-dimensional analysis* Scanning Microscopy Supplement 6:129-135.

F. C. A. Green et al. (1985) *A comparison of different focus functions for use in autofocus algorithms* Cytometry 6:81091.

S. Greenberg, D. Kogan (2005) *Structure-adaptive anisotropic filter applied to fingerprints* Optical Engineering 44(12):127004.

D. A. Grigg et al. (1992) *Probe characterization for scanning probe metrology* Utramicroscopy 42-44: 1616-1620.

W. Grimson (1981) From Images to Surfaces MIT Press, Cambridge, MA.

S. Grossberg, ed. (1988) Neural Computers and Natural Intelligence MIT Press, Cambridge, MA.

J. Gu, R. W. Ogilvie (2005) Virtual Microscopy and Virtual Slides in Teaching, Diagnosis and Research CRC Press, Boca Raton, FL.

P. Gualtieri, P. Coltelli (1992) *An automated system for the analysis of moving images* J. Comput. Assist. Microscopy 3(1):15-22.

M. Guarnera et al. (2010) *Adaptive color demosaicing and false color removal* J. Electronic Imaging 19(2): 031105.

H. J. G. Gundersen (1986) *Stereology of arbitrary particles* J. Microscopy 143:3-45.

H. J. Gundersen (2002) *The smooth fractionator* J. Microscopy 207:191-210.

H. J. G. Gundersen et al. (1978) *Distribution of membrane thickness determined by lineal analysis* J. Microscopy 113:27.

H. J. G. Gundersen, E. B. Jensen (1987) *The efficiency of systematic sampling in stereology and its prediction* J. Microscopy 147:229-263.

H. J. G. Gundersen et al. (1988) *Some new, simple and efficient stereological methods and their use in pathological research and diagnosis* Acta Pathologica, Microbiologica et Immunologica Scandinavica 96:857.

H. J. G. Gundersen et al. (1999) *The efficiency of systematic sampling in stereology — reconsidered* J. Microscopy 193(3):199-211.

F. Guo et al. (2010) *A novel multi-scale edge detection technique based on wavelet analysis with application in multiphase flows* Powder Technology 202:171-177.

D. -P. Hader, ed. (1992) Image Analysis in Biology, CRC Press, Boca Raton, FL.

A. J. Haines, J. S. Crampton (2000) *Improvements to the method of Fourier shape analysis as applied in morphometric studies* Paleontology 43(4):765-783.

J. Hajnal et al. (2001) Medical Image Registration, CRC Press, Boca Raton, FL.

K. J. Halford, K. Preston (1984) *3-D skeletonization of elongated solids* Computer Vision Graphics and Image Processing 27:78-91.

Y. S. Ham (1993) *Differential Absorption Cone-beam Microtomography* Ph. D. Thesis, North Carolina State University.

A. M. Hamed (2006) *Computation of the lateral and axial point spread functions in confocal imaging systems using binary amplitude mask* Pramana 66(6):1037-1048.

A. Hanbury, J. Serra (2002) *Mathematical morphology in the CIELAB space* Image Anal. Stereol 21:201-206.

D. J. Hand (1981) Discrimination and Classification, Wiley, New York.

R. Hanke (2003) *Microanalysis by volume computed tomography* G. I. T. Imaging and Microscopy 4:40-43.

R. M. Haralick (1978) *Statistical and structural approaches to texture* Proc 4th Intl Joint Conf Patt Recog, Kyoto, p. 45.

R. Haralick (1979) *Statistical and structural approaches to textures* Proc. IEEE 67:786-804.

R. M. Haralick, I. Dinstein (1975) *A spatial clustering procedure for multi-image data* Comput. Graph. Image Proc. 12:60-73.

R. M. Haralick, L. G. Shapiro (1988) *Segmentation and its place in machine vision* Scanning Microscopy Supplement 2:39-54.

R. M. Haralick, L. G. Shapiro (1992) Computer and Robot Vision I, Addison Wesley, Reading, MA.

R. M. Haralick et al. (1973) *Textural features for image classification* IEEE Trans. SMC 3:610-621.

J. A. Hartigan (1975) Clustering Algorithms John Wiley & Sons, New York.

R. V. L. Hartley (1942) *A more symmetrical Fourier analysis applied to transmission problems* Proc. IRE.

M. Hasegawa et al. (1996) *Calculation of fractal dimensions of machined surface profiles* Wear 192:40-45.

S. Haykin (1993) Neural Networks MacMillan, New York NY.

D. Hearn, M. P. Baker (1986) Computer Graphics, Prentice Hall, Englewood Cliffs, NJ.

J. Heath (2005) Dictionary of Microscopy Wiley, New York.

M. Heath et al. (1997) *A robust visual method for assessing the relative performance of edge-detection algorithms* IEEE Trans. Patt. Anal. Mach. Intell. PAMI-19 (12):1338-1359.

D. O. Hebb (1949) The Organization of Behaviour John Wiley, New York.

P. Heckbert (1982) *Color image quantization for frame buffer display* Computer Graphics 16(3):297-307.

D. Heeger, J. Bergen (1995) *Pyramid based texture analysis/synthesis* in Proc. SIGGRAPH, pp. 229-238.

R. Hegerl (1989) *Three-dimensional reconstruction from projections in electron microscopy* European Journal of Cell Biology 48 (Supplement 25):135-138.

H. Heijmans (1991) *Theoretical aspects of grey-level morphology* IEEE Trans PAMI 13(6):568-582.

H. Heijmans (1994) Morphological Image Operators Academic Press, New York.

C. Heipke (1992) *A global approach for least-squares image matching and surface reconstruction in object space* Photogrammetric Engineering and Remote Sensing 58(3):317-323.

J. Van Helden (1994) *CrestPathway algorithm* (personal communication).

L. Helfen et al. (2003) *Determination of structural properties of light materials* G. I. T. Imaging and Microscopy 4:55-57.

H. R. J. R. Van Helleputte et al. (1995) *Comparative study of 3D measurement techniques (SPM, SEM, TEM) for submicron structures* Microelectronic Engineering 27:547.

G. T. Herman (1980) Image Reconstruction from Projections—The Fundamentals of Computerized Tomography Academic Press, New York.

D. W. Hetzner (2008) *Comparing binary image analysis measurements—Euclidean geometry, centroids and corners* Microscopy Today July:10-15.

E. C. Hildreth (1983) *The detection of intensity changes by computer and biological vision systems* Comput. Vis. Graph. Image Proc. 22:1-27.

J. E. Hilliard, L. R. Lawson (2003) Stereology and Stochastic Geometry Soringer, New York.

K. Hoffmann et al. (2005) *Simple tool for the standardisation of confocal spectral imaging systems* GIT Imaging and Microscopy 3:18-19.

H. Hogan (2005) *Where's that picture?* Biophotonics International, July: 32-26.
S. G. Hoggar (2006) Mathematics of Digital Images: Creation, Compression, Restoration, Recognition Cambridge Univ. Press, Cambridge, UK.
T. J. Holmes et al. (1995) *Light microscopic images reconstructed by maximum likelihood deconvolution* in Handbook of Biological Confocal Microscopy (J. Pawley, ed.) Plenum Press, New York, 389-402.
T. J. Holmes, N. J. O'Connor (2000) *Blind deconvolution of 3D transmitted light brightfield micrographs* J Microscopy, 200(2):114-127.
G. C. Holst, T. S. Lomheim (2007) CMOS/CCD Sensors and Camera Systems SPIE Press, Bellingham, WA.
B. J. Holt, L. Hartwick (1994) *Visual image retrieval for applications in art and art history*, Proc. of Storage and Retrieval for Image and Video Databases II, SPIE, San Jose, CA.
L. Holzer et al. (2004) *Three-dimensional analysis of porous BaTiO3 ceramics using FIB nanotomography* J Microscopy 216(1):84-95.
L. Holzer et al. (2006) *FIB-nanotomography of particulate systems* J. Am. Ceram. Soc. 89(8):2577-2595.
L. Holzer et al. (2007) *Cryo-FIB-nanotomography for quantitative analysis of particle structures in cement suspensions* J. Microscopy 227(3):216-228.
B. K. P. Horn (1970) *Shape from Shading: A method for obtaining the shape of a smooth opaque object from one view* (AI Tech Report 79, Project MAC) Mass. Inst. Tech., Cambridge, MA.
B. K. P. Horn (1975) *Obtaining shape from shading information*, in Psychology of Computer Vision, P. H. Winston, ed., pp. 115-155, McGraw Hill, New York.
P. Hough (1962) *Method and means for recognizing complex patterns* U. S. Patent 3, 069, 654.
C. V. Howard, M. G. Reed (1998) Unbiased Stereology, Three Dimensional Measurements in Stereology BIOS Scientific Publishers, Oxford, UK.
C. V. Howard, M. G. Reed (2005) Unbiased Stereology: Advanced Methods Bios Scientific Publishers, Oxford, UK.
J. Hsieh (2003) Computed Tomography SPIE Press, Bellingham, WA.
J. Hsieh (2009) Computed Tomography: Principles, Design, Artifacts and Recent Advances SPIE Press, Bellingham, WA.
J. Huang et al. (1994) *A method for detecting correspondences in stereo pairs of electron micrographs of networks* Journal of Computer Assisted Microscopy 6(2):85-102.
T. S. Huang (1979) *A fast two dimensional median filtering algorithm* IEEE Trans ASSP-27:13-18.
D. H. Hubel (1988) Eye, Brain, and Vision Scientific American Library, W. H. Freeman, New York.
H. E. Hurst et al. (1965) Long Term Storage: An Experimental Study Constable, London.
C. Hwang et al. (1993) *Human visual system weighted progressive image transmission using lapped orthogonal transform classified vector quantization* Optical Engineering 32(7):1524-1530.
S. Inoue (1986) Video Microscopy Plenum Press, New York.
S. Inoué, K. R. Spring (1997) Video Microscopy - The Fundamentals (2nd Ed.) Plenum Press, New York.
V. K. Ivanov et al. (2007) *Fractal analysis of sea ice images* Sixth Int'l Kharkov Symposium on Physics and Engineering of Microwaves, 2:25-30, IEEE.
R. W. Ives et al. (2008) *Effects of image compression on iris recognition system performance* J. Electronic Imaging 17(1):011015.
H. Iwata, Y. Ukai (2002) *SHAPE: A computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors* Journal of Heredity 93:384-385.
G. H. Jacobs, J. Nathans (2009) *The evolution of primate color vision* Scientific American 300(4):56-63.
B. Jahne (1997) Practical Handbook on Image Processing for Scientific Applications CRC Press, Boca Raton, FL.
A. K. Jain (1989) Fundamentals of Digital Image Processing Prentice Hall, London.
A. Jalba et al. (2004) *Automatic segmentation of diatom images for classification* Microscopy Research and Technique 65:72-85.
M. James (1988) Pattern Recognition, Blackwell Scientific, London.
J. Jan (2005) Medical Image Processing, Reconstruction and Restoration, CRC Press, Boca Raton, FL.

J. R. Janesick (2001) Scientific Charge-Coupled Devices SPIE Press, Bellingham, WA.

G. Jayalaitha, R. Uthayakumar (2007) *Estimating the skin cancer using fractals* Proc. Int'l Conf. on Computational Intelligence v:306-311.

E. T. Jaynes (1985) *Where do we go from here?* in Maximum Entropy and Bayesian Methods in Inverse Problems (C. R. Smith, W. T. Grandy, ed.), Reidel Publishing Co., Dordrecht, Holland, 21-58.

K. F. Jarausch et al. (1996) *Silicon structures for in-situ characterization of atomic force microscopy probe geometry* J. Vac. Sci. Technol. B 14(6):3425.

J. P. Jernot (1982) Thése de Doctorat és Science, Université de Caen, France.

Y. Jiang et al. (2008) *Texture description based on multiresolution moments of image histograms* Optical Engineering 47(3):037005.

L. Jin et al. (2008) *Improved adaptive spatial distance-weighted median filter* J. Optical Engineering 46(3): 037003.

E. M. Johnson, J. J. Capowski (1985) *Principles of reconstruction and three-dimensional display of serial sections using a computer* in The Microcomputer in Cell and Neurobiology Research (R. R. Mize, ed.) Elsevier, New York, p. 249-263.

L. R. Johnson, A. K. Jain (1981) *An efficient two-dimensional FFT algorithm*, IEEE Trans PAMI-3:698-701.

Q. C. Johnson et al. (1986) *Micro-Tomography using Synchrotron Radiation* Lawrence Livermore National Laboratory Preprint UCRL-93538A. C. Kak, M. Slaney (1988) Principles of Computerized Tomographic Imaging IEEE Pub. PC-02071.

S. G. Johnson, M. Frigo (2007) *A modified split-radix FFT with fewer arithmetic operations* IEEE Trans. Signal Processing 55(1):111-119.

S. Joshi, M. I. Miller (1993) *Maximum a posteriori estimation with good roughness for 3D optiocal sectioning microscopy* Opt. Soc. Am. A. 10:1078-1985.

M. Kaes et al. (1987) *Snakes: Active contour models*, J. Computer Vision 1:321-331.

A. C. Kak, M. Slaney (1987, 2001) Principles of Computerized Tomographic Imaging SIAM, Philadelphia, PA.

H. R. Kang (1997) Color Technology for Electronic Imaging Devices SPIE Optical Engineering Press, Bellingham, WA.

H. R. Kang (1999) Digital Color Halftoning SPIE Optical Engineering Press, Bellingham, WA.

L. -W. Kang, J-J. Leon (2005) *Fast indexing and searching strategies for feature-based image database systems* J. Electronic Imaging 14:013019.

J. N. Kapur et al. (1985) *A new method for grey-level picture thresholding using the entropy of the histogram* Computer Vision, Graphics and Image Processing 29:273-285.

N. Karssemeijer et al. (eds.) Digital Mammography Kluwer Academic, New York.

M. Kass et al. (1987) *Snakes: Active Contour Models* International Journal of Computer Vision 1(4):321-331.

A. E. Kayaalp, R. C. Jain (1987) *Using SEM stereo to extract semiconductor wafer pattern topography* Proc. SPIE 775:18-26.

B. H. Kaye et al. (1983) *A study of physical significance of three-dimensional signature waveforms* Proc. Fineparticle Characterization Conference.

B. H. Kaye (1986) *Image analysis procedures for characterizing the fractal dimension of fine particles* Proc. Particle Technol. Conf., Nürnberg.

B. H. Kaye (1989) A Random Walk Through Fractal Dimensions VCH Publishers, Weinheim, Germany.

A. Keating (1993) Visualization of a bat's ear, Duke Univ., private communication.

D. J. Keller (1991) *Reconstruction of STM and AFM images distorted by finite-size tips* Surf. Sci. 253:353-364.

D. J. Keller, F. S. Franke (1993) *Envelope reconstruction of probe microscope images* Surf. Sci. 294:409-419.

G. M. P. Van Kempen et al. (1997) *A quantitative comparison of image restoration methods in confocal microscopy* J. Microscopy 185(3):354-365.

M. R. Khadivi (1990) *Iterated function system in generating fractal fractures models scaling* in Disordered Materials: Fractal Structures and Dynamics, p. 49-51, Materials Research Society, Pittsburgh, PA.

A. Kikuchi et al. (2002) *Fractal tumor growth of ovarian cancer: sonographic evaluation* Gynecologic Oncology

87(3):295-302.

D. -W. Kim, K. -S. Hong (2006) *Real-time mosaic using sequential graph* J. Electronic Imaging 15(2):023005.

E. Kim et al. (2003) *A high resolution multispectral imaging systems for small satellites* Acta Astronautica 52:813-818.

S. Kim et al. (1988) *A new chain-coding algorithm for binary images using run-length codes* Computer Vision, Graphics & Image Processing 41:114-128.

S. H. Kim, J. P. Allebach (2005) *Optimal unsharp mask for image sharpening and noise removal* J. Electronic Imaging 14(2):023005.

C. Kimme et al. (1975) *Finding circles by an array of accumulators* Communication of the ACM 18(2).

J. H. Kinney et al. (1989) *X-ray microtomography on beamline X at SSRL* Rev. Sci. Instrum. 60(7):2471-4.

J. H. Kinney et al. (1990) *Nondestructive imaging of materials microstructures using X-ray tomographic microscopy* Proc. MRS Symposium on Tomographic Imaging, Boston, MA.

R. Kirsch (1971) *Computer determination of the constituent structure of biological images* Comput. Biomed. Res. 4:315-328.

J. Kittler et al. (1985) *Threshold selection based on a simple image statistic* Computer Vision, Graphics and Image Processing 30:125-147.

V. Kober et al. (2001) *Nonlinear filters with spatially connected neighborhood* Opt. Eng. 40(6):971-983.

L. P. Kok (1990) 100 Problems of my Wife, and Their Solution In Theoretical Stereology Coulomb Press, Leyden, The Netherlands.

H. Kotera, H. Wang (2005) *Multiscale image sharpening adaptive to edge profile* J. Electronic Imaging 14: 013002.

W. Kou (1995) Digital Image Compression: Algorithms and Standards Springer, New York.

A. Kriete, ed. (1992) Visualization in Biomedical Microscopies: 3D Imaging and Computer Applications VCH Publishers, Weinheim.

C. Kubel et al. (2005) *Recent advances in electron tomography: TEM and HAADF-STEM tomography for materials science and semiconductor applications* Microscopy and Microanalysis 11:378-400.

L. Kubinova, J. Janacek (2001) *Confocal microscope and stereology: estimating volume, number, surface area and length by virtual test probes applied to three-dimensional images* Microscopy Reseach and Technique 53:425-435.

F. P. Kuhl, C. R. Giardina (1982) *Elliptic Fourier features of a closed contour* Computer Graphics and Image Processing 18:236-258.

K. J. Kurzydlowski, B. Ralph (1995) The Quantitative Description of the Microstructure of Materials, CRC Press, Boca Raton, FL.

M. Kuwahara et al. (1976) *Processing of RI-angiocardiographic images* in Digital Processing of Biomedical Images (K. Preston & M. Onoe, ed.), pp. 187-202, Plenum, New York.

R. L. Lagendijk, J. Biemond (1991) Iterative Identification and Restoration of Images Kluwer Academic, Boston, MA.

A. Laine, W. Huda (2000) *Enhancement by multiscale nonlinear operators* in Handbook of Medical Imaging, I. Bankmann ed., Academic Press, London, UK.

L. Lam et al. (1992) *Thinning methodologies - a comprehensive survey* IEEE Trans PAMI 14:868-885.

L. Landini et al., ed. (2006) Advanced Image Processing in Magnetic Resonance Imaging CRC Press, Boca Raton, FL.

L. Landweber (1951) *An iterative formula for Fredholm integral equations of the first kind* Amer. J. Math., 73(3):615-624.

C. Lantejoul, S. Beucher (1981) *On the use of the geodesic metric in image analysis* J. Microscopy 121:39.

R. S. Ledley et al. (1990) *Fundamentals of true-color image processing* Proc. Intl. Conf. Patt. Recog. 1:791-795.

C. P. Lee (2006) *Robust Image Segmentation Using Active Contours: Level Set Approaches* Ph. D. Thesis, N. C. State Univ. Dept. of Electrical and Computer Engineering.

D. L. Lee, A. T. Winslow (1993) *Performance of three image-quality metrics in ink-jet printing of plain papers* J Electronic Imaging 2(3):174-184.

S. U. Lee et al. (1990) *A comparative performance study of several global thresholding techniques for segmentation* Computer Vision, Graphics and Image Processing 52:171-190.

A. Lehman et al. (2005) JMP for Basic Univariate and Multivariate Statistics SAS Press, Cary, NC.

A. Leonard et al. (2007) *Non-destructive characterization of deer antlers by X-ray microtomography coupled with image analysis* J. Microscopy 225(3):258-263.

O. Leroux et al. (2009) *A new preparation method to study fresh plant structures with X-ray computed tomography* J. Microscopy 233(1):1-4.

Z. Les, M. Les (2005) *Shape understanding system: understanding of the complex object* J. Electronic Imaging 14(2) 023015.

P. E. Lestrel, ed. (1997) Fourier Descriptors and Their Applications in Biology, Cambridge Univ. Press, Cambridge.

J. Y. Lettvin et al. (1959) *What the frog's eye tells the frog's brain* Proc. Inst. Rad. Eng. 47(11):1940-1951.

S. Levialdi (1972) *On shrinking binary picture patterns* Communications of the ACM 15(1):7-10.

H. Li, M. Novak, R. Forchheimer (1993) *Fractal-based image sequence compression scheme* Optical Engineering 32(7):1588-1595.

B. Lichtenbelt et al. (1998) Introduction to Volume Rendering Prentice Hall, Saddle River, NJ.

W. Lin et al. (1994) *Data corrections for three-dimensional light microscopy stereo pair reconstruction* Journal of Computer Assisted Microscopy 6(3):113-128.

M. Lineberry (1982) *Image segmentation by edge tracing* Applications of Digital Image Processing IV, vol 359.

X. Y. Liu et al. (2007) *Dynamic evaluation of autofocusing for automated microscopic analysis of blood smear and pap smear* J. Microscopy 27(1):15-23.

S. Lobregt et al. (1980) *Three-dimensional skeletonization: principle and algorithm* IEEE Trans. PAMI-2:75-77.

S. Lovejoy (1982) *Area-perimeter relation for rain and cloud areas* Science 216:185-187.

R. L. Luck et al. (1993) *Morphological Classification System and Method* U. S. Patent 5, 257, 182.

L. B. Lucy (1974) *An iterative technique for the rectification of observed distributions* Astronomical Journal 79(6):745-754.

M. Luessi et al. (2009) *Framework for efficient optimal multilevel image thresholding* J. Electronic Imaging 18(1):013004.

Z. P. Luo, J. H. Koo (2007) *Quantifying the dispersion of mixture microstructures* J. Microscopy 225(2): 118-125.

E. Mach (1906) *Über den Einfluss räumlich und zeitlich variierender Lichtreize auf die Gesichtswarhrnehmung* S. -B. Akad. Wiss. Wien, Math. -Nat. Kl. 115:633-648.

J. B. MacQueen (1967) *Some methods for the classification and analysis of multivariate observations* Proc. Of the 5th Berkeley Symposium on Mathematical Statistics and Probability, 1, 281-297.

E. Mainsah et al. (2001) *Surface measurement and characterization* in Metrology and Properties of Engineering Surfaces (E. Mainsah, ed.) Kluwer Academic Publisher, London, 1-42.

S. G. Mallat (1989) *A theory for multiresolution signal decomposition: the wavelet representation* IEEE Trans. Patt. Anal. Mach. Intell. PAMI-11(7):674-693.

T. Malzbender et al. (2001) *Polynomial texture maps* Computer Graphics, Proceedings of Siggraph 2001, pp. 519-528.

C. A. Mandarim-de-Lacerda (2003) *Stereological tools in biomedical research* Anais da Academia Brasileira de Ciencias 75(4):469-486.

B. B. Mandelbrot (1967) *How long is the coast of Britain? Statistical self-similarity and fractional dimension* Science 155:636-638.

B. B. Mandelbrot (1982) The Fractal Geometry of Nature W. H. Freeman, San Francisco.

B. B. Mandelbrot, D. E. Passoja, A. J. Paullay (1984) *Fractal character of fracture surfaces of metals* Nature 308:721.

E. Manders et al. (1993) *Measurement of co-localization of objects in dual-color confocal images* J. Microscopy 169:375-382.

R. Mann et al. (2001) *Augmented reality visualization of fluid mixing in stirred chemical reactors using electrical resistance tomography* J. Electronic Imaging 10(3):620-629.

P. Markiewicz, M. C. Goh (1994) *Atomic force microscopy probe tip visualization and improvement of images using a simple deconvolution procedure* Langmuir 10:5-7.

P. Markiewicz, M. C. Goh (1995) *Atomic force microscope tip deconvolution using calibration arrays* Rev. Sci. Instrum. 66:3186-3190.

D. Marr (1982) Vision W. H. Freeman, San Francisco.

D. Marr, E. Hildreth (1980) *Theory of edge detection* Proc. R. Soc. Lond. B207:187-217.

D. Marr, T. Poggio (1976) *Cooperative computation of stereo disparity*, Science, 194:283-287.

M. De Marsicoi et al. (1997) *Indexing pictorial documents by their content: a survey of current techniques* Image and Vision Computing 15:119-141.

M. Martin-Landrove et al. (2007) *Fractal analysis of tumoral lesions in brain*, in Engineering in Medicine and Biology, p. 1306-1309, IEEE.

C. G. Masi (2005) *Dynamic structured light measures shapes* Vision Systems Design, March 2005, p. 15-20.

G. A. Mastin (1985) *Adaptive filters for digital image noise smoothing: An evaluation* Comput. Vis. Graph. Image Proc. 31:102-121.

D. Matacara-Hernandez (2002) Color Vision and Colorimetry SPIE Press, Bellingham, WA.

Y. A. T. Mattfeldt (2005) *Explorative statistical analysis of planar point processes in microscopy* J. Microscopy 220(3):131-139.

N. A. Matthews (2008) *Aerial and Close-Range Photogrammetic Technology: Providing Resource Documentation, Interpretation and Preservation*, Technical Note 428, U. S. Dept. of the Interior Bureau of Land Management, National Operations Center, Denver, CO.

A. D. McAulay, J. Wang, J. Li (1993) *Optical wavelet transform classifier with positive real Fourier transform wavelets* Optical Engineering 32(6):1333-1339.

C. S. McCamy et al. (1976) *A color-rendition chart* Journal of Applied Photographic Engineering 11(3):95-99.

J. J. McCann (2004) *Capturing a black cat in shade: past and present of Retinex color appearance models* J. Elect. Imaging 13(1) 36-47.

W. S. McCulloch, W. Pitts (1943) *A logical calculus of the ideas immanent in nervous activity* Bull. Math. Biophys. 5:115.

J. J. Mecholsky, D. E. Passoja (1985) *Fractals and brittle fracture* in Fractal Aspects of Materials p. 117-119, Materials Research Society, Pittsburgh, PA.

J. J. Mecholsky et al. (1986) *Crack propagation in brittle materials as a fractal process* in Fractal Aspects of Materials II Materials Research Society, Pittsburgh, PA.

J. J. Mecholsky et al. (1989) *Quantitative analysis of brittle fracture surfaces using fractal geometry* J. Am. Ceram. Soc 72:60.

G. Medioni, R. Nevatia (1985) *Segment-based stereo matching* Computer Vision Graphics and Image Proc. 31:2-18.

F. Melgani (2006) *Robust image binarization with ensembles of thresholding algorithms* J. Electronic Imaging 15(2):023010.

P. Miché et al. (2005) *Passive 3-D shape recovery of unknown objects using cooperative polarimetric and radiometric stereo vision processes* Optical Engineering 44(2):027005.

D. L. Milgram (1975) *Computer methods for creating photomosaics* IEEE Trans C-24:1113-1119.

D. L. Milgram, M. Herman (1979) *Clustering edge values for threshold selection* Comput. Graph. Image Proc. 10:272-280.

M. Minsky, S. Papert (1969) Perceptrons: An Introduction to Computational Geometry MIT Press, Cambridge, MA.

M. W. Mitchell, D. A. Bonnell (1990) *Quantitative topographic analysis of fractal surfaces by scanning tunneling microscopy* Journal of Materials Research 5(10):2244-2254.

S. Mitra, T. Acharya (2003) Data Mining. Multimedia, Soft Computing and Bioinformatics Wiley, Hoboken, NJ.

J. Modersitzki (2004) Numerical Methods for Image Registration Oxford Univ. Press, Oxford, UK.

J. R. Monck et al. (1992) *Thin-section ratiometric Ca2+ images obtained by optical sectioning of Fura-2 loaded mast cells* J. Cell Biol. 116:745-759.

C. Montagne et al. (2006) *Adaptive color quantization using the baker's transformation* J. Electronic Imaging 15(2):023015.

H. P. Moravec (1977) *Towards automatic visual obstacle avoidance* Proc. 5th IJCAI:584.

S. Morillas et al. (2007) *New adaptive vector filter using fuzzy metrics* J. Electronic Imaging 16(3):033007.

K. Mossberg, M. Ericsson (1990) *Detection of doubly stained fluorescent specimens using confocal microscopy* J. Microscopy158:215-224.

R. B. Mott (1995) *Position-tagged spectrometry, a new approach for EDS spectrum imaging* Proc. Microscopy and Microanalysis, p. 595, Jones & Begall, NY.

J. C. Mott-Smith (1970) *Medial axis transformations* in Picture Processing and Psychopictorics (B. S. Lipkin, A. Rosenfeld, eds.) Academic Press, New York.

P. Mouton (2002) Principles and Practices of Unbiased Stereology: An Introduction for Bioscientists Johns Hopkins Univ. Press, Baltimore, MD.

P. R. Mouton et al. (2002) *Stereological length estimation using spherical probes* J. Microscopy 206(1):54-64.

M. Mudge, T. Malzbender (2006) *New reflection transformation imaging methods for rock art and multipleviewpoint display* IEEE Symposium on Virtual Analytics Science and Technology.

H. R. Myler, A. R. Weeks (1993) Pocket Handbook of Image Processing Algorithms in C Prentice Hall, Englewood Cliffs, NJ.

A. Myronenko, X. Song (2009) *Global active contour-based image segmentation via probability alignment* Computer Vision Pattern Recognition pp. 2798-2804.

J. Nakamura (ed.) (2006) Image Sensors and Signal Processing for Digital Still Cameras CRC Press, Boca Raton, FL.

K. S. Nathan, J. C. Curlander (1990) *Reducing speckle in one-look SAR images* NASA Tech Briefs, Feb:70.

F. Natterer (2001) The Mathematics of Computerized Tomography SIAM, Philadelphia, PA.

F. Natterer, F. Wubbeling (2001) Mathematical Methods in Image Reconstruction SIAM, Philadelphia, PA.

B. Neal et al. (1998) *A super-resolution approach to perimeter measurement* J. Computer Assisted Microscopy, 10(1):11-22.

B. Neal, J. C. Russ (2004) *Principal Components Analysis Of Multispectral Images* Microscopy Today 12(5):36.

A. J. Nederbracht et al. (2004) *Image calibration, filtering and processing*, p. 35-58 in E. Francus (ed.) Image Analysis, Sediments and Paleoenvironments Springer, Dordrecht, The Netherlands.

C. V. Negoita, D. A. Ralescu (1987) Simulation, Knowledge-based Computing, and Fuzzy Statistics Van Nostrand Reinhold, New York.

C. V. Negoita, D. A. Ralescu (1975) Applications of Fuzzy Sets to Systems Analysis Halsted Press, New York.

A. Neubeck, L. VanGool (2006) *Efficient non-maximum suppression* Proc. 18th Int'l Conf. Pattern Recognition, 850-855.

R. Nevatia, K. Babu (1980) *Linear feature extraction and description* Computer Graphics and Image Processing, vol. 13.

W. Niblack (ed.) (1993) *Storage and Retrieval for Image and Video Databases* SPIE Proceedings Vol. 1908.

A. Nicoulin et al. (1993) *Subband image coding using jointly localized filter banks and entropy coding based on vector quantization* Optical Engineering 32(7):1430-1450.

A. Nieminen et al. (1987) *A new class of detail-preserving filters for image processing* IEEE Trans. Patt. Anal. Mach. Intell. PAMI-9:74-90.

N. Nikolaidis, I. Pitas (2001) 3-D Image Processing Algorithms Wiley, New York.

S. De Nooijer et al. (2008) *Sphere size distributions from finite thickness sections: a forward approach employing a genetic algorithm* J. Microscopy 231(2):257-264.

J. F. O'Callaghan (1974) *Computing the Perceptual Boundaries of Dot Patterns* Computer Graphics and Image Processing 3(2):141-162.

D. Ochoa et al. (2010) *Automatic identification of C. elegans in population images by shape energy features*, J. Microscopy 238(2):173-184.

J. Ohser, K. Schladitz (2009) 3D Images of Materials Microstructures — Processing and Analysis Wiley- VCH, Weinheim.

K. Oistämö, Y. Neuvo (1990) *Vector median operations for color image processing* Nonlinear Image Processing (E. J. Delp. ed.) SPIE Proc. 1247:2-12.

C. K. Olsson (1993) *Image Processing Methods in Materials Science*, Ph. D. Thesis, Technical Univ. of Denmark, Lyngby, Denmark.

J. D. Ortiz, S. O'Connell (2004) *Toward a non-linear grayscale calibration method for legacy photographic collections* in E.

Francus (ed.) Image Analysis, Sediments and Paleoenvironments Springer, Dordrecht, The Netherlands, p. 125-141.

N. Otsu (1979) *A threshold selection method from gray-level histograms* IEEE Trans SMC-9:62-69; 377-393.

D. R. Overby, M. Johnson (2005) *Studies on depth-of-field effects in microscopy supported by numerical simulations* J. Microscopy 220(3):176-189.

I. Overington (1976) Vision and Acquisition Pentech Press, London.

I. Overington (1992) Computer Vision: A Unified Biologically Inspired Approach Elsevier, Amsterdam.

Y. A. Ozkaya et al. (2005) *Digital image processing and illumination techniques for yarn characterization* J. Electronic Imaging 14(2):023001.

J. Padfield et al. (2005) *Polynomial texture mapping: A new tool for examining the surface of paintings* ICOM Committee for Conservation, 1:504-510.

J. Paluszynski, W. Slowko (2009) *Measurements of the surface microroughness with the scanning electron microscope* J. Microscopy 233(1):10-17.

M. Pancorbo et al. (1991) *New filtering techniques to restore scanning tunneling microscopy images* Surface Science 251/252:418-423.

D. P. Panda, A. Rosenfeld (1978) *Image segmentation by pixel classification in (gray level, edge value) space* IEEE Trans. Comput. 27:875-879.

Y. -H. Pao (1989) Adaptive Pattern Recognition and Neural Networks Addison Wesley, Reading, MA.

J. Park et al. (2005) *Fast disparity estimation algorithm using the property of stereo matching* Optical Engineering 44(6):060501.

K. Parker et al. (2005) *Color in medical imaging* Biophotonics International 1:44-48.

J. R. Parker (1997) Algorithms for Image Processing and Computer Vision John Wiley & Sons, New York.

T. Pavlidis (1977) Structural Pattern Recognition Springer Verlag, New York.

T. Pavlidis (1980) *A thinning algorithm for discrete binary images* Computer Graphics and Image Processing 13:142-157.

T. Pavlidis (1982) Algorithms for Graphics and Image Processing Computer Science Press, Rockville, MD.

L. D. Peachey, J. P. Heath (1989) *Reconstruction from stereo and multiple electron microscope images of thick sections of embedded biological specimens using computer graphics methods* J. Microscopy, 153:193-204.

D. M. Pearsall (1978) *Phytolith analysis of archaeological soils: evidence for maize cultivation in formative Ecuador* Science 199(4325):177-178.

S. Peleg et al. (1984) *Multiple resolution texture analysis and classification* IEEE Trans Patt Anal Mach Intell PAMI-6:518.

A. P. Pentland (1983) *Fractal-based description of natural scenes* IEEE Trans PAMI-6:661.

A. P. Pentland, ed. (1986) From Pixels to Predicates Ablex, Norwood, NJ.

A. P. Pentland et al. (1994) *Photobook: Content-based manipulation of image databases* Proc. of Storage and Retrieval Image and Video Databases II, SPIE, San Jose, CA; Int'l Journal of Computer Vision 18: 233-254.

E. Persoon, K. -S. Fu (1977) *Shape descrimination using Fourier descriptors* IEEE Trans. Sys. Man Cyber. SMC-7:170-179.

J. L. Pfalz (1976) *Surface networks* Geographical Analysis 8(2):77-93.

G. Piazzesi (1973) *Photogrammetry with the scanning electron microscope* J. Phys. E Sci Instrum. 6:392-396.

J. Piper (2008) *Use of software to enhance depth of field and improve focus in photomicrography* Microscopy and Analysis May:15-19.

D. R. Piperno (1984) *A comparison and differentiation of phytoliths from maize (Zea mays L.) and wild grasses: Use of morphological criteria* American Antiquity 49(2):361-383.

I. Pitas (2000) Digital Image Processing Algorithms and Applications Wiley, New York.

E. Ponz et al. (2006) *Measuring surface topography with scanning electron microscopy* Microscopy and Microanalysis 12:170-177.

W. K. Pratt (1991) Digital Image Processing 2nd edition Wiley, New York.

T. Prettyman et al. (1991) *On the performance of a combined transmission and scattering approach to industrial computed*

tomography Advances in X-ray Analysis vol. 35, Plenum Press, New York.

J. M. S. Prewitt, M. L. Mendelsohn (1966) *The analysis of cell images*, Ann. N. Y. Acad. Sci 128:1035-1053.

I. Przerada, A. Bochinek (1990) *Microfractographical aspects of fracture toughness in microalloyed steel*, in Stereology in Materials Science Polish Soc. for Stereology, Krakow.

J. Puzicha et al. (1998) *On Spatial Quantization of Color Images* Tech. Rept IAI-TR-98-1 Univ. of Bonn.

L. Quam, M. J. Hannah (1974) *Stanford automated photogrammetry research* AIM-254, Stanford AI Lab.

C. F. Quate (1994) *The AFM as a tool for surface imaging* Surf. Sci. (Netherlands) 299-300, 980-95.

M. Rabbani, ed. (1992) Selected Papers on Image Coding and Compression SPIE Press, Bellingham. WA.

J. Radon (1917) *Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten* Berlin Sächsische Akad. Wissen. 29:262-279.

Z. Rahman et al. (2004) *Retinex processing for automatic image enhancement* J. Elect. Imaging 13(1):100-110.

R. Ramanath (2000) *Interpolation methods for the Bayer color array*, M. S. Thesis, N. C. State Univ. Dept of Electrical Eng. (www4. ncsu. edu/~rramana/Research/MastersThesis. pdf).

R. Rangayyan (2005) Biomedical Image Analysis CRC Press, Boca Raton, FL.

M. Raspanti et al. (2005) *A vision-based 3D reconstruction technique for scanning electron microscopy: Direct comparison with atomic force microscopy* Microscopy Research and Technique 67, p. 1-7.

K. K. Ray, G. Mandala (1992) *Study of correlation between fractal dimension and impact energy in high strength low alloy steel* Acta Metallurgica et Materialia 40(3):463.

B. S. Reddy, B. N. Chatterji, (1996) *An FFT-based technique for translation, rotation, and scale-invariant image registration* IEEE Trans. Image Processing 5(8):1266-1271.

M. G. Reed et al. (1997) *Confocal imaging and second-order stereological analysis of a liquid foam* J. Microscopy 185(3):313-320.

M. G. Reed, C. V. Howard (1997) *Edge corrected estimates of the nearest neighbor function for three-dimensional point patterns* J. Microscopy 186(2): 177-184.

M. G. Reed, C. V. Howard (1998) Unbiased Stereology Bios Scientific Publisher, Oxford.

A. A. Reeves *Optimized Fast Hartley Transform with Applications in Image Processing* Thesis, Dartmouth University, March 1990.

R. G. Reeves, ed. (1975) Manual of Remote Sensing American Society of Photogrammetry, Falls Church, CA.

K. Rehm et al. (1994) *Display of merged multimodality brain images using interleaved pixels with independent color scales*, J. Nucl. Med. 35:1815-1821.

E. Reinhard et al. (2002) *Photographic tone reproduction for digital images* ACM Transactions on Graphics 21(3) (Proceedings of Siggraph 2002).

G. Reiss et al. (1990) *Scanning tunneling microscopy on rough surfaces: Tip-shape-limited resolution* J. Appl. Phys. 67(3):1156-1159.

I. Rezanaka, R. Eschbach, eds. (1996) Recent Progress in Ink Jet Technologies Soc. Imaging Science & Tech, Springfield, VA.

H. Rheingold (1991) Virtual Reality Summit, New York.

W. H. Richardson (1972) *Bayesian-based iterative method of image restoration* J. Opt. Soc. Am. 62(1):55-59.

J. P. Rigaut (1988) *Automated image segmentation by mathematical morphology and fractal geometry* J. Microscopy 150:21-30.

G. X. Ritter, J. N. Wilson (2001) Handbook of Computer Vision Algorithms in Image Algebra, 2nd ed. CRC Press, Boca Raton, FL.

M. Rizon et al. (2005) *Object detection using circular Hough transform* Amer. J. Applied Sciences 2(12).

L. G. Roberts (1965) *Machine perception of three-dimensional solids* in Optical and Electro-Optical Information Processing (J. T. Tippett, ed.) MIT Press, Cambridge, MA.

L. G. Roberts (1982) *Recognition of three-dimensional objects* in The Handbook of Artificial Intelligence vol III (P. Cohen, E. Figenbaum, ed.) Kaufmann, Los Gatos, CA.

I. Rock (1984) Perception W. H Freeman, New York.

F. J. Rohlf (1990) *Morphometrics* Annu. Rev. Ecol. Syst. 21:299-316.

F. J. Rohlf, J. W. Archie (1984) *A comparison of Fourier methods for the description of wing shape in mosquitoes (Diptera: culicidae)* Syst. Zool. 33:302-317.

D. W. Rolston (1988) Principles of Artificial Intelligence and Expert System Development McGraw Hill, New York.

B. G. Rosen, R. J. Crafoord, eds. (1997) Metrology and Properties of Engineering Surfaces Chalmers Univ., Göteborg, Sweden.

F. Rosenblatt (1958) *The perceptron: A probabilistic model for information organization and storage in the brain* Psych. Rev. 65:358-408.

A. Rosenfeld, J. Pfaltz (1968) *Distance functions on digital pictures*. Patt. Recog. 1(1):33-61.

A. Rosenfeld, A. C. Kak (1982) Digital Picture Processing vol 1 & 2, Academic Press, New York.

F. S. Rosenthal, Z. A. Begum (2005) *Image-based determination of chord lengths in air-dried lungs* J Microscopy 219(3):160-166.

M. Rossner, K. M. Yamada (2004) *What's in a picture? The temptation of image manipulation* J. Cell Biology 166(1):11-15.

I. Rovner (1971) *Potential of opal phytoliths for use in paleoecological reconstruction* Quaternary Research 1(3):345-59.

Y. Rui et al. (1999) *Image retrieval: Current techniques, promising directions and open issues* Journal of Visual Communication and Image Representation 10(1):39-62.

J. Ruiz-Alzola et al. (2005) *Landmark-based registration of medical image data* in L Costaridou (ed.) Medical Image Analysis Methods CRC Press, Boca Raton, FL.

D. E. Rumelhart et al. (1986) *Learning representations by back-propagating errors* Nature 323:533-536.

J. C. Russ (1984) *Implementing a new skeletonizing method* J. Microscopy 136:RP7.

J. C. Russ (1986) Practical Stereology Plenum Press, New York.

J. C. Russ (1988) *Differential absorption three-dimensional microtomography* Trans Amer. Nucl. Soc 56(3):14.

J. C. Russ (1990a) Computer Assisted Microscopy, Plenum Press, New York.

J. C. Russ (1990b) *Surface characterization: Fractal dimensions, Hurst coefficients and frequency transforms* J. Comput. Assist. Microsc. 2(3):161-184.

J. C. Russ (1990c) *Processing images with a local Hurst operator to reveal textural differences*, J. Comput. Assist. Microsc. 2(4):249-257.

J. C. Russ (1991) *Multiband thresholding of images* J. Comput. Assist. Microsc. 3(2):77-96.

J. C. Russ (1993) *JPEG image compression and image analysis* J. Comput. Assist. Microsc. 5(3):237-244.

J. C. Russ (1993) *Method and application for ANDing features in binary images* J. Comput. Assist. Microsc. 5(4):265-272.

J. C. Russ (1994) Fractal Surfaces Plenum Press, New York.

J. C. Russ (1995a) *Computer-assisted manual stereology* J. Comput. Assist. Microsc. 7(1):35-46.

J. C. Russ (1995b) *Median filtering in color space* J. Comput. Assist. Microsc. 7(2):83-90.

J. C. Russ (1995c) *Thresholding Images* J. Comput. Assist. Microsc. 7(3):41-164.

J. C. Russ (1995d) *Designing kernels for image filtering* J. Comput. Assist. Microsc. 7(4):179-190.

J. C. Russ (1995e) *Optimal greyscale images* J. Comput. Assist. Microsc. 7(4):221-234.

J. C. Russ (1995f) *Segmenting touching hollow features* J. Comput. Assist. Microsc. 7(4):253-261.

J. C. Russ (1997) *Fractal dimension measurement of engineering surfaces* in 7th Int'l Conf. on Metrology and Properties of Engineering Surfaces (B. G. Rosen, R. J. Crafoord, eds.), Chalmers Univ., Göteborg, Sweden, 170-174.

J. C. Russ (2001a) Forensic uses of digital microscopy CRC Press, Boca Raton FL.

J. C. Russ (2001b) *Fractal geometry in engineering metrology* in Metrology and Properties of Engineering Surfaces (E. Mainsah et al., eds.) Kluwer Academic Publishers, London, 43-82.

J. C. Russ (2002) Forensic Uses of Digital Imaging CRC Press, Boca Raton, FL.

J. C. Russ (2004) Image Analysis of Food Microstructure CRC Press, Boca Raton, FL.

J. C. Russ (2009) The Image Processing Cookbook Amazon Createspace, Scotts Valley, CA.

J. C. Russ et al. (1989) *Application of the Hough transform to electron diffraction patterns* J. Comput. Assist. Microsc. 1(1):3-77.

J. C. Russ. R. T. Dehoff (2001) Practical Stereology, 2nd Edition, Plenum Press.

J. C. Russ, I. Rovner (1987) *Stereological verification of* Zea *Phytolith taxonomy* Phytolitharien Newsletter, 4(3): 10.

J. C. Russ, J. C. Russ (1988a) *Automatic discrimination of features in grey scale images* J. Microscopy 148: 263-277.

J. C. Russ, J. C. Russ (1988b) *Improved implementation of a convex segmentation algorithm* Acta Stereologica 7:33-40.

J. C. Russ, J. C. Russ (1989a) *Uses of the Euclidean distance map for the measurement of features in images* J. Comput. Assist. Microsc. 1(4):343.

J. C. Russ, J. C. Russ (1989b) *Topological measurements on skeletonized three-dimensional networks* J. Comput. Assist. Microsc. 1:131-150.

J. C. Russ, J. C. Russ (2007) Introduction to Image Processing and Analysis CRC Press, Boca Raton, FL.

J. C. Russ et al. (1989) *Direct 3-D pore location measurement in alumina* J. Microscopy 155(2):RP1-2.

P. Russell, D. Batchelor (2001) *SEM and AFM: Complementary techniques for surface investigations* Microscopy and Analysis, July 2001:5-8.

M. R. Rutenberg et al. (2001) *Automated Cytological Specimen Classification System and Method* U. S. Patent 6, 327, 377.

F. F. Sabins, Jr. (1987) Remote Sensing: Principles and Interpretation (2nd Edition) W. H. Freeman, New York.

F. Sacerdotti et al. (2002) *Closed regions: a proposal for spacing parameters for areal surface measurements* Meas. Sci. Technol. 13:556-564.

P. K. Sahoo et al. (1988) *A survey of thresholding techniques* Computer Vision Graphics and Image Processing 41:233-260.

E. Sanchez, L. A. Zadeh, ed. (1987) Approximate Reasoning in Intelligent System Decision and Control. Oxford Press, New York.

J. Sanchez, M. P. Canton (1999) Space Image Processing CRC Press, Boca Raton, FL.

K. Sandau, J. Ohser (2007) *The chord length transform and the segmentation of crossing fibers* J. Microscopy 226(1):43-53.

L. M. Sander (1986) *Fractal growth processes* Nature 322:789-793.

B. Sapoval (1991) *Fractal electrodes, fractal membranes, and fractal catalysts*, in Fractals and Disordered Systems Springer Verlag, Berlin.

L. J. Sartor, A. R. Weeks (2001) *Morphological operations on color images* Journal of Electronic Imaging 10(2):548-559.

S. Saunders (1991) *Magellan: the Geologic Exploration of Venus* Engineering and Science, Spring 1991:15-27.

A. Savitsky, M. J. E. Golay (1964) *Smoothing and differentiation of data by simplified least squares procedures* Anal. Chem. 36:1627-1639.

R. J. Schalkoff (1991) Pattern Recognition: Statistical, Syntactical and Neural Approaches. Wiley, New York.

D. J. Schneberk et al. (1991) *Multiple energy techniques in industrial computerized tomograhy*, in Review of Progress in Quantitative Nondestructive Evaluation (D. O. Thompson, D. E. Chimenti, eds.), Plenum Press, New York.

R. Schneider, W. Weil (2008) Stochastic and Integral Geometry. Springer, New York.

H. P. Schwartz, K. C. Shane (1969) *Measurement of particle shape by Fourier analysis* Sedimentology 13: 213-231.

H. Schwarz, H. E. Exner (1980) *Implementation of the concept of fractal dimensions on a semi-automatic image analyzer* Powder Technology 27:107.

H. Schwarz, H. E. Exner (1983) *The characterization of the arrangement of feature centroids in planes and volumes* J. Microscopy 129:155.

C. Sciammarella et al. (2005) *High accuracy contouring using projection moiré* Optical Engineering 44(9): 093605.

P. J. Scott (1995) *Recent advances in areal characterization* IX Intern. Oberflächenkolloq, Tech. Univ. Chemnitz- Zwickau, 151-158.

P. J. Scott (1997) *Foundations of topological characterization of surface texture* in 7th Int'l Conf. on Metrology and Properties of Engineering Surfaces (B. G. Rosen, R. J. Crafoord, eds.), Chalmers Univ., Göteborg, Sweden, 162-169.

J. P. Secilla, N. Garcia (1987) *Template location in noisy pictures* Signal Process 14:347-361.

J. Serra (1982) Image Analysis and Mathematical Morphology Academic Press, London.

M. Seul et al. (2000) Practical Algorithms for Image Analysis Cambridge Univ. Press, Cambridge.

M. Sezgin, B. Sankur (2004) *Survey over image thresholding techniques and quantitative performance evaluation* J. Electronic Imaging 13(1):146-165.

M. Shao et al. (2005) *Partition-based interpolation for color filter array demosaicking and super-resolution reconstruction* Optical Engineering 44(10):107003.

L. G. Shapiro, R. M. Haralick (1985) *A metric for comparing relational descriptions* IEEE Trans. Pattern Anal. and Mach. Intell. PAI-7(1):90-94.

G. Sharma (ed.) 2003) Digital Color Imaging Handbook CRC Press, Boca Raton, FL.

G. Sharma (2005) *Imaging arithmetic*: *Physics U Math > Physics + Math* SPIE Electronic Imaging Tech. Group Newsletter 15(2):1-10.

L. A. Shepp, B. F. Logan (1974) *The Fourier reconsruction of a head section* IEEE Trans NS-21:21-43.

A. Shih et al. (2001) *Fast algorithm for X-ray cone-beam microtomography* Microscopy and Microanalysis, 7:13-23.

F. Y. Shih (2009) Image Processing and Mathematical Morphology CRC Press, Boca Raton, FL.

S. O. Shim et al. (2010) *Accurate shape from focus based on focus adjustment in optical microscopy* Microscopy Research and Technique 72:362-370.

D. Shuman (2005) *Computerized image analysis software for measuring indents by AFM* Microscopy and Analysis, May 2005, 15-17.

S. A. Sirr, J. R. Waddle (1999) *The utility of computed tomography in the detection of internal damage and repair and the determination of authenticity of high quality bowed stringed instruments* RadioGraphics, May.

J. Skilling (1986) *Theory of maximum entropy image reconstruction,* in Maximum Entropy and Bayesian Methods in Applied Statistics, Proc. 4th Max Entropy Workshop, Univ of Calgary, 1984 (J. H. Justice, ed.) pp. 156-178, Cambridge Univ. Press, Cambridge.

P. E. Slatter (1987) Building Expert Systems: Cognitive Emulation Halsted Press, New York.

B. D. Smith (1990) *Cone-beam tomography*: *recent advances and a tutorial review* SPIE Optical Engineering 29:5.

P. W. Smith, M. D. Elstrom (2001) *Stereo-based registration of range and projective imagery for data fusion and visualization* Optical Engineering 40(3):352-361.

R. F. Smith (1990) Microscopy and Photomicrography CRC Press, Boca Raton, FL.

B. Smolka et al. (2001) *On the reduction of impulse noise in multichannel image processing* Opt. Eng. 40(6): 902-908.

D. L. Snyder et al. (1992) *Deblurring subject to nonnegative constraints* IEEE Trans. Sig. Proc. 40:1143-1150.

I. Sobel (1970) *Camera models and machine perception* AIM-21, Stanford Artificial Intelligence Lab, Palo Alto.

R. Sobol (2004) *Improving the Retinex algorithm for rendering wide dynamic range photographs* J. Elect. Imaging 13(1) 65-74.

P. Soille (1999) Morphological Image Analysis Springer Verlag, Berlin.

Y. Solihin et al. (1996) *A comparison of thresholding techniques for handwriting and text extraction* 4th Int'l Conf. on Control, Automation, Robotics, and Vision (ICCARV), Singapore, 1408-1412.

M. Sonka et al. (1999) Image Processing, Analysis and Machine Vision, 2nd ed. Brooks Cole, Pacific Grove, CA.

P. G. Spetsieris et al. (1995) *Interactive visualization of coregistered tomographic images* Biomedical Visualization (BioMedVis '95), p. 58.

S. Srinivasan et al. (1990) *Fractal analysis of erosion surfaces* Journal of Materials Research 5(11):2616-2619.

J. A. Stark, W. J. Fitzgerald (1996) *An alternative algorithm for adaptive histogram equalization* Comp. Vis. Graph. Image Proc. 56(2):180-185.

M. Stefik (1995) Introduction to Knowledge Systems Morgan Kaufmann, San Francisco.

D. C. Sterio (1984) *The unbiased estimation of number and sizes of arbitrary particles using the disector* J. Microscopy 134:127-136.

P. L. Stewart, R. M. Burnett Seminars in Virology (1991) cited in C. J. Mathias *Visualization techniques augment lab research into structure of adenovirus* Scientific Computing & Automation 7(6):51-56.

M. C. Stone et al. (1988) *Color gamut mapping and the printing of digital color images* ACM Trans. Graphics 7(4):249-292.

J. A. Storer (1992) Image and Text Compression Kluwer Academic Publishers, New York.

K. J. Stout et al. (1993) *The development of methods for the characterization of roughness in three dimensions*, publication EUR 15178 EN of the Commission of the European Communities, Univ. of Birmingham, Edgbaston, England.

T. Strang (1996) Wavelets and Filter Banks Wellesley Coll, Wellesley, MA.

R. G. Summers et al. (1991) *The use of confocal microscopy and stereocon reconstructions in the analysis of sea urchin emryonic cell division* J. Electron Microscope. Tech. 18:24-30.

C. Sun, P. Vallotton (2010) *Fast linear feature detection using multiple directional non-maximum suppression* J. Microscopy 234(2):147-157.

H. Q. Sun, Y. J. Luo (2009) *Adaptive watershed segmentation of binary particle image* J. Microscopy 233(2): 326-330.

Y. Sun et al. (2004) *Autofocusing in Computer Microscope: Selecting the Optimal Focus Algorithm* Microscopy Research and Technique 65, 139-149.

I. Sutherland (1965) *The ultimate display* Proceedings of IFIP 65, vol. 2:506-508, 582-583.

A. Swaminathan et al. (2007) *Non-intrusive component forensics of visual sensors using output images* IEEE Transactions on Information Forensics and Security 2(1):91-106.

R. E. Swing (1997) An Introduction to Microdensitometry SPIE Press, Bellingham, WA.

P. J. Tadrous (2009) *A method of PSF generation for 3D brightfield deconvolution* J. Microscopy 227(2): 192-199.

H. Talbot et al. (2000) *Image analysis of insulation mineral fibers* J. Microscopy 200(3):251-258.

D. Tamburrino et al. (2010) *Two pixels, full color: Image sensor demosaicing and characterization* J. Electronic Imaaging 19(2):021103.

L. Tang et al. (2005) *Novel dense matching algorithm with Voronoi decomposition of images* Optical Engineering 44(10):107201.

P. Thevenaz, M. Unser (2006) *User-friendly semiautomated assembly of accurate image mosaics in microscopy* Microscopy Research and Technique 70:135-146.

T. R. Thomas (1999) Rough Surfaces, 2nd edition Imperial College Press, London.

M. M. Thompson, ed. (1966) Manual of Photogrammetry American Society of Photogrammetry, Falls Church, VA.

M. von Tiedemann et al. (2006) *Image adaptive point spread function estimation and deconvolution for in vivo confocal microscopy* Microsc. Research & Technique 69:10-20.

J. T. Tou, R. C. Gonzalez (1981) Pattern Recognition Principles Addison Wesley, Reading, MA.

J. Trussell (1979) *Comments on "picture thresholding using an iterative selection method"* IEEE Trans SMC 9:311.

T. -M. Tu et al. (2001) *Efficient intensity-hue-saturation-based image fusion with saturation compensation* Opt. Eng. 40(5):720-728.

E. R. Tufte (1990) Envisioning Information Graphics Press, Cheshire CT.

E. R. Tufte (1997) Visual Explanations: Images and Quantities, Evidence and Narrative Graphics Press, Cheshire, CT.

E. R. Tufte (2001) The Visual Display of Quantitative Information 2nd Edition, Graphics Press, Cheshire, CT.

J. N. Turner et al. (1991) *Confocal microscopy and three-dimensional reconstruction of electrophysiologically identified neurons in thick brain slices* J. Electron Microscope Tech. 8:11-23.

F. Tushabe, M. H. F. Wilkinson (2008) *Content-based image retrieval using combined 2D attribute pattern spectra* in Advances in Multilingual and Multimodal Information Retrieval, Lecture Notes in Computer Science, Springer, Vol. 5152.

P. C. Twiss et al. (1969) *Morphological classification of grass phytoliths* Soil Science Society of America Proceedings 33(1):109-115.

R. M. Udre, N. Vizireanu (2007) *Iterative generalization of morphological skeleton* J. Electronic Imaging Letters 16(1):010501.

S. E. Umbaugh (1998) Computer Vision and Image Processing Prentice Hall, Saddle River, NJ.

E. E. Underwood (1970) Quantitative Stereology Addison Wesley, Reading, MA.

E. E. Underwood, K. Banerji (1986) *Fractals in fractography* Materials Science and Engineering 80:1.

S. Uppal et al. (2010) *Morphological fractal analysis of shape in cancer cells* Microscopy & Microanalysis 16: 472-477.

A. Ursani et al. (2008) *Texture features based on local Fourier histogram: self-compensation against rotation* J. Electronic Imaging Letters 17(3):030503.

W. Vanderlinde, J. Caron (2008) *Blind deconvolution of SEM images* Scanning 30(3):268-269.

J. G. Verly, R. L. Delanoy (1993) *Some principles and applications of adaptive mathematical morphology for range imagery* Optical Engineering 32(12):3295-3306.

H. Verschueren et al. (1993) *Methods for computer assisted analysis of lymphoid cell shape and motility, including Fourier analysis of cell outlines* J. Immunol. Meth. 163:99-113.

T. Vicsek (1992) Fractal Growth Phenomena 2nd ed. World Scientific, Singapore.

J. S. Villarrubia (1994) *Morphological estimation of tip geometry for scanned probe microscopy* Surf. Sci 321:287-300.

J. S. Villarrubia (1996) *Scanned probe microscope tip characterization without calibrated tip characterizers* J. Vac. Sci. Technol. B14:1518-1521.

J. S. Villarubbia (1997) *Algorithms for Scanned Probe Microscopy Image Simulation, Surface Reconstruction and Tip Estimation* J. Res. Nat'l. Inst. Stand. Tech. 102(4):425.

C. Vonesch, M. Unser (2008) *A fast thresholded Landweber algorithm for wavelet-regularized multidimensional deconvolution* IEEE Trans. on Image Processing 17(4):539-549.

G. K. Wallace (1991) *The JPEG still picture compression standard* Comm. ACM 34:30-44.

G. Wang et al. (1991) *Scanning cone-beam reconstruction algorithms for X-ray microtomography* SPIE Scanning Microscope Instrumentation 1556:99.

G. Wang, M. Vannier (2001) *Micro-CT Scanners for Biomedical Applications*, Advanced Imaging, July 2001: 18-27.

H. Wang et al. (2008) *Accurate length measurement of multiple cotton fibers* J. Electronic Imaging 17(3):031110.

R. J. Wall et al. (1974) *Analysis of image histograms* Proc. 2nd Joint Int'l Conf. Patt. Recog. IEEE 74CH-0885-4C:341-344.

J. R. Walters (1988) Crafting Knowledge-based Systems: Expert Systems Made Easy Wiley, New York.

Z. Wang (1990) Principles of Photogrammetry (With Remote Sensing) Press of Wuhan Technical University of Surveying and Mapping, Beijing.

J. Wasen, R. Warren (1990) Catalogue of Stereological Characteristics of Selected Solid Bodies, Volume 1: Polyhedrons Chalmers Univ., Göteborg, Sweden.

J. Wasen, et al. (1996) Catalogue of Stereological Characteristics of Selected Solid Bodies, Volume 2: Hexagonal Prisms Chalmers Univ., Göteborg, Sweden.

J. Wasen, et al. (1996) Catalogue of Stereological Characteristics of Selected Solid Bodies, Volume 3: Ellipsoids Chalmers Univ., Göteborg, Sweden.

J. Wasen et al. (1996) Catalogue of Stereological Characteristics of Selected Solid Bodies, Volume 4: Cubospheres Chalmers Univ., Göteborg, Sweden.

A. R. Weeks (1996) Fundamentals of Electronic Image Processing SPIE Press, Bellingham, WA.

D. Wehbi et al. (1992) *Perturbation dimension for describing rough surfaces*, Int'l Jnl of Machine Tools and Manufacturing 32:211.

E. R. Weibel (1979) Stereological Methods vol. I & II, Academic Press, London.

B. Weiss (2006) *Fast median and bilateral filtering* ACM Trans. Graph. 25(3):519-526.

S. Welstead (1999) Fractal and Wavelet Image Compression Techniques SPIE Press, Bellingham, WA.

A. Wen, C. Lu (1993) *Hybrid vector quantization* Optical Engineering 32(7):1496-1502.

J. S. Werner et al. (2007) *Illusory color and the brain* Scientific American 296(3):90-95.

J. West et al. (1997) *Comparison and Evaluation of retrospective intermodality registration techniques* Journal of Computer Assisted Tomography, 21:54-566.

J. S. Weszka (1978) *A survey of threshold selection techniques* Comput. Graph. Image. Proc. 7:259-265.

J. Weszka et al. (1976) *A comparative study of texture measures for terrain classification* IEEE Trans SMC- 6:269-285.

J. S. Weszka, A. Rosenfeld (1979) *Histogram modification for threshold selection* IEEE Trans SMC-9:38-52.

D. J. Whitehouse (1994) Precision - the Handbook of Surface Metrology Institute of Physics Publishing, Bristol.

H. K. Wickramasinghe (1991) *Scanned probes old and new* AIP Conf. Proc. (USA), 9-22.

S. D. Wicksell (1925) *The corpuscle problem* Biometrica 17(25):84 and 18(26):152.

B. Willis et al. (1993) *Iterative, constrained 3D image reconstruction of transmitted light bright field micrographs based on maximum likelihood estimation* J. Microscopy 169:347-361.

R. Wilson, M. Spann (1988) Image Segmentation and Uncertainty Wiley, New York.

G. Winstanley (1987) Program Design for Knowledge-based Systems Halsted Press, New York.

R. J. Woodham (1978) *Photometric stereo: a reflectance map technique for determining surface orientation from image intensity* Proc. SPIE 155:136-143.

G. Wolf (1991) *Usage of global information and a priori knowledge for object isolation* Proc. 8th Int'l. Congr. for Stereology, Irvine, CA, 56.

K. W. Wong (1980) *Basic mathematics of photogrammetry*, in Manual of Photogrammetry, 4th Edition, American Society of Photogrammetry, Falls Church, Va, 57-58.

B. P. Wrobel (1991) *Least-squares methods for surface reconstruction from images* ISPRS J. Photogrammetry and Remote Sensing, 46:67-84.

H. -S. Wu et al. (2005) *Segmentation of intestinal gland images with iterative region growing* Journal of Microscopy 220(3):190-204.

S. Wu, A. Gersho (1993) *Lapped vector quantization of images* Optical Engineering 32(7):1489-1495.

X. Wu et al. (2005) *Novel fractal image-encoding algorithm based on a full binary tree searchless iterated function system* Optical Engineering 44(10):107002.

Z. Q. Wu et al. (2005) *Adaptive contrast enhancement based on highly overlapped interpolation* J. Electronic Imaging 14(3):033006.

X. Xiang et al. (2006) *Low complexity and high efficiency image compression algorithm* J. Electronic Imaging 15(2):023017.

H. Xie et al. (2007) *Construction and evaluation of a wavelet-based focus measure for microscopy imaging* Microscopy Research and Technique 70:987-995.

R. R. Yager (1979) *On the measures of fuzziness and negation, part 1: membership in the unit interval* Int. Journal of Gen. Sys. 5:221-229.

G. J. Yang, T. S. Huang (1981) *The effect of median filtering on edge location estimation* Comput. Graph. Image Proc. 15:224-245.

N. Yang, J. Boselli, I. Sinclair (2001) *Simulation and quantitative assessment of homogeneous and inhomogeneous particle distributions in particulate metal matrix composites* J. Microscopy 201(2):189-200.

Y. Yakimovsky (1976) *Boundary and object detection in real world images* J. Assoc. Comput. Mach. 23: 599-618.

H. Yoo et al. (2006) *Measurement and restoration of the point spread function of fluorescence confocal microscopy* J. Microscopy 221(3):172-176.

B. -W. Yoon, W. -J. Song (2007) *Image contrast enhancement based on the generalized histogram* J. Electronic Imaging 16(3):033005.

R. W. Young, N. G. Kingsbury (1993) *Video compression using lapped transforms for motion estimation/compensation and coding* Optical Engineering 32(7):1451-1463.

T. York (2001) *Status of electrical tomography in industrial applications* J. Electronic Imaging 10(3):608-619.

A. Yoshitaka, T. Ichikawa (1999) *A survey on content-based retrieval for multimedia databases*, IEEE Trans. Knowledge and Data Engineering 11(1):81-93.

Q. Yu et al. (2004) *A novel contour-based 3D terrain matching algorithm using wavelet transform* Pattern Recogn. Lett. 25(1):87-99.

Z. Yuan, J. Li (2008) *Robust image restoration for rotary motion blur based on frequency analysis* Optical Engineering 47:097004.

L. A. Zadeh (1965) *Fuzzy sets* Information and Control 8:338-353.

C. T. Zahn, R. Z. Roskies (1972) *Fourier descriptors for plane closed curves* IEEE Trans. Comput. C-21:269-281.

G. Zhai et al. (2008) *Unified deblocking for discrete cosine transfer compressed images* J. Electronic Imaging 17(4): 043021.

G. Zhang (2009) *Novel template matching method with sub-pixel accuracy based on correlation and Fourier- Mellin transform* Optical Engineering 48:057001.

J. Zhang, Q. Zhang (2007) *Noniterative blind image restoration based on estimation of a significant class of point-spread functions* Optical Engineering 46(7):077005.

X. Zhou, E. Dorrer (1994) *An automatic image matching algorithm based on wavelet decomposition* ISPRS International Archives of Photogrammetry and Remote Sensing 30(3/2):951-960.

Q. Z. Zhu, Z. M. Zhang (2005) *Correlation of angle-resolved light scattering with the microfacet orientation of rough silicon surfaces* Optical Engineering 44(7) 073601.

H. -J. Zimmermann (1987) Fuzzy Sets, Decision Making and Expert Systems Kluwer Academic Publishers, Boston, MA.

T. G. Zimmerman et al. (1987) *A hand gesture interface device* Proceedings of CHI+GI '87, p. 189-192.

O. Zinchuk, V. Zinchuk (2006) *Dynamics of cellular responses studied by quantitative colocalization analysis* Microscopy and Analysis 20:S13-S15.

V. Zinchuk et al. (2007) *Quantitative colocalization analysis of multicolor confocal immunofluoresscence microscopy images: Pushing pixels to exlpore biological phenomena* Acta Histochem Cytochem 40:101-111.

V. Zinchuk, O. Zinchuk (2008) *Quantitative colocalization analysis of confocal fluorescence microscopy images* Current Protocols in Cell Biology Unit 4. 19, Wiley, New York.

A. Zizzari (2004) Methods on Tumor Recognition and Planning Target Prediction for the Radiotherapy of Cancer Shaker-Verlag, Aachen, Germany.

S. Zucker (1976) *Region growing: Childhood and adolescence* Comput. Graph. Image Proc. 5:382-399.

O. Zuniga, R. M. Haralick (1983) *Corner detection using the facet model* Comput. Vision Pattern Recog. IEEE, 30-37.

J. M. Zwier et al. (2004) *Image calibration in fluorescence microscopy* J. Microscopy 216(1):15-24.